



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 184052

TO: Manjunath N Rao
Location: REM-2A01/2C70
Art Unit: 1652
Tuesday, April 11, 2006
Case Serial Number: 09/211691

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

184052
STIC-Biotech/ChemLib

From: Rao, Manjunath N.
Sent: Tuesday, April 04, 2006 8:40 AM
To: STIC-Biotech/ChemLib
Subject: Sequenc esearch request for 09/211691

RECEIVED
APR - 4 2006
STIC/BIOTECH/CHMLIB
(STIC)

From: Manjunath N. Rao
Art Unit 1652, Room 2A01
Mail Box in Room 2C70
Phone: 272-0939

Date: 4-4-06

Please search the following as soon as possible for application with serial number
09/211691

1. SEQ ID NO: 1, 3, 4 against all commercial nucleic acid sequence databases, issued patents/published applications nucleic acid sequence database and pending application nucleic acid sequence database. Please provide a print of results
2. SEQ ID NO: 2 against all commercial amino acid sequence databases, issued patents/published applications amino acid sequence database and pending application amino acid sequence database. Please provide a print of results.

If you have any questions please call me at the above phone number.

Thanks

4/4/2006

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:52 ; Search time 5660.11 Seconds
(without alignments)
8315.447 Million cell updates/sec

Title: US-09-211-691-1
Perfect score: 828
Sequence: 1 atgcaaacacagcttcag.....tcattgtgccttcacataa 828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_srb:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hlg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	828	100.0	3613	1	NM025839	U25839 Neisseria m
2	826.4	99.8	110000	1	AE002098_20	Continuation (21 o
3	826.4	99.8	349980	6	AX044034	AX044034 Sequence
4	808.8	97.7	4827	1	AY039684	AY039684 Neisseria
5	800.8	96.7	3335	1	AF470659	AF470659 Neisseria
6	791.2	95.6	2959	1	AF240672	AF240672 Neisseria
7	788	95.2	3369	1	AF470660	AF470660 Neisseria
8	786.4	95.0	3367	1	AF470658	AF470658 Neisseria
9	782	94.4	5216	1	AY134877	AY134877 Neisseria
10	775.6	93.7	3289	1	AF470663	AF470663 Neisseria
11	775.6	93.7	3289	1	AF470662	AF470662 Neisseria
12	753.2	91.0	3383	1	AF470655	AF470655 Neisseria
13	750	90.6	3259	1	AF470665	AF470665 Neisseria
14	748.4	90.4	3066	1	AY134878	AY134878 Neisseria
15	746.8	89.2	349061	1	NMA222491	AL162753 Neisseria
16	740	89.4	3266	1	AF470661	AF470661 Neisseria
17	736.8	89.0	3136	1	AF470656	AF470656 Neisseria
18	736.8	89.0	3287	1	AF470657	AF470657 Neisseria

19	735.2	88.8	4517	1	AF355193	AF355193 Neisseria
20	726	87.7	5859	1	NCU14554	U14554 Neisseria g
21	726	87.7	5859	6	AR025350	AR025350 Sequence
22	726	87.7	5859	6	AR025351	AR025351 Sequence
23	726	87.7	5859	6	AR111460	AR111460 Sequence
24	726	87.7	5859	6	I24768	I24768 Sequence 1
25	726	87.7	5859	6	I24769	I24769 Sequence 7
26	726	87.7	5859	6	AR183924	AR183924 Sequence
27	726	87.7	5859	6	AR183925	AR183925 Sequence
28	726	87.7	5859	6	AR207874	AR207874 Sequence
29	726	87.7	5859	6	AR567686	AR567686 Sequence
30	726	87.7	5859	6	AR567687	AR567687 Sequence
31	721.2	87.1	53922	1	AB004969_21	Continuation (22 o
32	710	85.7	4883	1	AF313394	AF313394 Neisseria
33	708.4	85.6	5262	1	AY134876	AY134876 Neisseria
34	706	85.3	4862	1	AF121135	AF121135 Neisseria
35	600.2	72.5	3851	1	AY039686	AY039686 Neisseria
36	549	66.3	3851	1	AY039685	AY039685 Neisseria
37	531.4	64.2	2343	1	AF355194	AF355194 Neisseria
38	523.4	63.2	2243	1	AF470664	AF470664 Neisseria
39	484.6	58.5	843	1	AF208063	AF208063 Neisseria
40	459.6	55.5	3712	1	NM065788	U65788 Neisseria m
41	426	51.4	840	6	AX565941	AX565941 Sequence
42	115.6	14.0	10827	1	AE006155	AE006155 Pasteurel
43	103.2	12.5	3998	1	AF224466	AF224466 Haemophil
44	103.2	12.5	30191	1	AE017152	AE017152 Haemophil
45	94.2	11.4	807	1	AY091470	AY091470 Haemophil

ALIGNMENTS

RESULT 1	NM025839	3613 bp	DNA	linear	BCT 18-MAY-1996
LOCUS	NM025839				
DEFINITION	Neisseria meningitidis glycy1 tRNA synthetase (glys) gene, partial cds, and glycosyl transferase (lgta), (lgta), (lgta) genes,				
ACCESSION	U25839				
VERSION	U25839.1	GI:973183			
KEYWORDS	Neisseria meningitidis				
SOURCE	Neisseria meningitidis				
ORGANISM	Neisseria meningitidis				
REFERENCE	Neisseria meningitidis				
AUTHORS	Jennings, M.P., Hood, D.W., Peak, I.R., Virji, M. and Moxon, E.R.				
TITLE	Molecular analysis of a locus for the biosynthesis and phase-variable expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in Neisseria meningitidis				
JOURNAL	Mol. Microbiol. 18 (4), 729-740 (1995)				
PUBMED	8817494				
REFERENCE	2 (bases 1 to 3613)				
AUTHORS	Jennings, M.P.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-APR-1995) Michael P. Jennings, Department of Pediatrics, University of Oxford, Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, Oxfordshire OX3, 9DU, United Kingdom				
FEATURES	location/Qualifiers				
source	1..3613				
gene	/organism="Neisseria meningitidis"				
CDS	/mol_type="genomic DNA"				
	/strain="MC58"				
	/db_xref="taxon:487"				
	1..264				
	/gene="glys"				
	<1..264				
	/gene="glys"				
	/codon_start=1				
	/transl_table=11				
	/product="glycy1 tRNA synthetase"				
	/protein_id="AAC44083.1"				

```
/db_xref="GI:973184"
/translation="SILQDEKALFAAQGIQPKIAAVAGNFQTLSELAVKPO
VAFPDGVVWMEADAVKONRLNRLAEQNNAAVADIALGE"
gene
318..1329
/gene="1gtA"
318..321
/gene="1gtA"
328..1329
/gene="1gtA"
/function="Involved in the biosynthesis of the terminal
lacto-N-neotetraose LPS structure"
/codon_start=1
/evidence=experimental
/translation="glycosyl transferase"
/product="glycosyl transferase"
/protein_id="AAC44084.1"
/db_xref="GI:973185"
/translation="MOPLVSLICAVNVEKPAOSIAAVNQTENLDLIVDSDT
GLTAAQROEDDRIILAOBRNSGLIPSLNIGDLAKSGGGSEYIARTPDIDIA
PMWIKIVGEMKRSIIAMGAMLEVLSEKDNRLAHHEKXIKWPTREDIADP
PFGNPINHTMTMRSSVIDGRLRTNEDNADBDYQWIDSKRLATYPAVKR
LHANGVSKTISIRHSLAQGIQKTAANDLQSGMFKTPDSLEYQIAVAVELLEKH
LPEDFELARRLYOCFKRTDTLPAGAWLDFADGRMRLFTLRQYFGLHRLKNR"
582..595
/misc_feature
/gene="1gtA"
/notes="encodes homopolymeric tract of 14 guanosine
residues which control the phase variable expression of
the 1gtA gene via a slipped-strand mispairing mechanism"
1361..2198
/gene="1gtB"
1361..1364
/gene="1gtB"
1371..2198
/gene="1gtB"
/function="Involved in the biosynthesis of the terminal
lacto-N-neotetraose LPS structure"
/codon_start=1
/evidence=experimental
/translation="glycosyl transferase"
/product="glycosyl transferase"
/protein_id="AAC44085.1"
/db_xref="GI:973186"
/translation="MGNHVISLASAERRAHADTFGRHGLPQFPFDALMPSERLEQA
MAELVGLAKOHLSEVEKACPMASHAVLMKQALDGLPYVAVFLGKDAEKFLADTW
LEBRLQKDSAFVRLLETWFAKVIYRDKLVNTENSFLBSEHNGTAGYIISREAMR
FLERFPAVLPAVADKAVDMMPFTYDFDKGMFVYQVNALCTQELHAKFLSKNMLG
SPLEKROERHRSRLKVMFLKRALGFKREKKKMRBQRALEKAYGSRVTSFK"
2337..3177
/gene="1gtB"
2337..2340
/gene="1gtB"
2347..3177
/gene="1gtB"
/function="Involved in the biosynthesis of the terminal
lacto-N-neotetraose LPS structure"
/codon_start=1
/evidence=experimental
/translation="glycosyl transferase"
/product="glycosyl transferase"
/protein_id="AAC44086.1"
/db_xref="GI:973187"
/translation="MGNHVISLASAERRAHADTFGRHGLPQFPFDALMPSERLENR
MAELVGLAKOHLSEVEKACPMASHAVLMKQALDGLPYVAVFLGKDAEKFLADTW
LEBRLQKDSAFVRLLETWFAKVIYRDKLVNTENSFLBSEHNGTAGYIISREAMR
FLERFPAVLPAVADKAVDMMPFTYDFDKGMFVYQVNALCTQELHAKFLSKNMLG
SPLEKROERHRSRLKVMFLKRALGFKREKKKMRBQRALEKAYGSRVTSFK"
repeat_region 3297..3369
repeat_region 3538..3606
ORIGIN
Query Match 100.0%; Score 828; DB 1; Length 3613;
Best Local Similarity 100.0%; Pred. No. 9.1e-18;
```

```
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTTATGAGTTAGCTTCCGCCGAGAAACGAGGCGCACATTTGCCAT 60
Db 1371 ATGCAAAACCAAGTTATGAGTTAGCTTCCGCCGAGAAACGAGGCGCACATTTGCCAT 1430
QY 61 ACCTTGGGAGGACGAGCATCCGTTTCAGTTTTTCAGCGACATGATCCGCTGAAAGG 120
Db 1431 ACCTTGGGAGGACGAGCATCCGTTTCAGTTTTTCAGCGACATGATCCGCTGAAAGG 1490
QY 121 CTGGAACGAGCAATGGCGAACTGCTCCCGGCTTGTGGGCAACCCCTATTGAGCGGA 180
Db 1491 CTGGAACGAGCAATGGCGAACTGCTCCCGGCTTGTGGGCAACCCCTATTGAGCGGA 1550
QY 181 GTGAAAAAGCCCTGTTATGAGCGACGCGTATTTGTGAAGCAGGCAATTGACGAGGT 240
Db 1551 GTGAAAAAGCCCTGTTATGAGCGACGCGTATTTGTGAAGCAGGCAATTGACGAGGT 1610
QY 241 CTGCGGTATATCAACGTTATTTGAGAGCAGTTTTTACTGGCGAAGGTGAGAAAAATTC 300
Db 1611 CTGCGGTATATCAACGTTATTTGAGAGCAGTTTTTACTGGCGAAGGTGAGAAAAATTC 1670
QY 301 CTTCGCGAAGACGCTTGCTGCAAGAACGCTTTTGAACCGGATACCGCTTATCGTCCGC 360
Db 1671 CTTCGCGAAGACGCTTGCTGCAAGAACGCTTTTGAACCGGATACCGCTTATCGTCCGC 1730
QY 361 TTGAAAACGATGTTATGACAGTCTGACCTCGCCCTCGGCGGAGATTACTGGCGG 420
Db 1731 TTGAAAACGATGTTATGACAGTCTGACCTCGCCCTCGGCGGAGATTACTGGCGG 1790
QY 421 CGCGCTTTCGCTGTTGGAAGCGAACACATGAGGAGGACGAGGCGTATATCTCCGA 480
Db 1791 CGCGCTTTCGCTGTTGGAAGCGAACACATGAGGAGGACGAGGCGTATATCTCCGA 1850
QY 481 AAAGCATGCGGTTTCTTGACAGAGTTTCCGCGCTGCGCCGGAAGGCTGACCCC 540
Db 1851 AAAGCATGCGGTTTCTTGACAGAGTTTCCGCGCTGCGCCGGAAGGCTGACCCC 1910
QY 541 GTGCACTGATGATGTTGAGGATTTTTCAGAGGAAAGAAATGCGGTTTGCAGCTC 600
Db 1911 GTGCACTGATGATGTTGAGGATTTTTCAGAGGAAAGAAATGCGGTTTGCAGCTC 1970
QY 601 AATCCGCGCTTGTGGCGCCAGAGCTGATATGCAAGTTTCAACGACCAAGCGCA 660
Db 1971 AATCCGCGCTTGTGGCGCCAGAGCTGATATGCAAGTTTCAACGACCAAGCGCA 2030
QY 661 TTGGGCAAGCTGATGGAACAGACCGCTCTGGAACCGCAACGCAAGGCGGATTC 720
Db 2031 TTGGGCAAGCTGATGGAACAGACCGCTCTGGAACCGCAACGCAAGGCGGATTC 2090
QY 721 CCCGCAACATTCATCAACACCGGCTGATCGGCGCTTGACCAAAATCAGCAGGAAAG 780
Db 2091 CCCGCAACATTCATCAACACCGGCTGATCGGCGCTTGACCAAAATCAGCAGGAAAG 2150
QY 781 GAAAAACGCGCGCAAGGCGGAAACAGTTGATTTGCTTTCCAAATA 828
Db 2151 GAAAAACGCGCGCAAGGCGGAAACAGTTGATTTGCTTTCCAAATA 2198
RESULT 2
AB002098_20/c
Sequence split into 23 fragments LOCUS AB002098 Accession AB002098
WPCOMMENT
Fragment Name Begin End
AB002098_00 1 11000
AB002098_01 10001 21000
AB002098_02 20001 31000
AB002098_03 30001 41000
AB002098_04 40001 51000
AB002098_05 50001 61000
AB002098_06 60001 71000
AB002098_07 70001 81000
AB002098_08 80001 91000
```

AE002098_09	900001	1010000
AE002098_10	100001	1110000
AE002098_11	110001	1210000
AE002098_12	120001	1310000
AE002098_13	130001	1410000
AE002098_14	140001	1510000
AE002098_15	150001	1610000
AE002098_16	160001	1710000
AE002098_17	170001	1810000
AE002098_18	180001	1910000
AE002098_19	190001	2010000
AE002098_20	200001	2110000
AE002098_21	210001	2210000
AE002098_22	220001	2272360

Query Match	99.8%;	Score	826.4;	DB	1;	Length	110000;
Best Local Similarity	99.9%;	Pred.	No. 2.5e-186;				
Matches	827;	Conservative	0;	Mismatches	1;	Indels	0;
				Gaps	0;		
QY	1	ATGCAAAACCA	CGTTATCAGCTT	AGCTTCGCGCAGAA	CGCAGGCGCACATTCGCAT	60	
Db	26697	ATGCAAAACCA	CGTTATCAGCTT	AGCTTCGCGCAGAA	CGCAGGCGCACATTCGCAT	26638	
QY	61	ACCTTCGCGCAGG	CAACGCGATCCCGTTTCAGTTTTCAGTCCGCTCTGAAAG	120			
Db	26637	ACCTTCGCGCAGG	CAACGCGATCCCGTTTCAGTTTTCAGTCCGCTCTGAAAG	26578			
QY	121	CTGGAACAGGCAAT	AGGCGGAACCTGTC	CCGCGCTGGCGCAGACCCATTATTGACGGGA	180		
Db	26577	CTGGAACAGGCAAT	AGGCGGAACCTGTC	CCGCGCTGGCGCAGACCCATTATTGACGGGA	26518		
QY	181	GTGGAABAAAG	CTTCCTTTATGAGC	CACGCGTATGTGGAAGCAGGCATTGACGAAGGT	240		
Db	26517	GTGGAABAAAG	CTTCCTTTATGAGC	CACGCGTATGTGGAAGCAGGCATTGACGAAGGT	26458		
QY	241	CTGCGCATATATAC	ACGATTTTGGAGAGAGAGTTTACTCGGCGCAAGGTGAGAAAATTTC	300			
Db	26457	CTGCGCATATATAC	ACGATTTTGGAGAGAGAGTTTACTCGGCGCAAGGTGAGAAAATTTC	26398			
QY	301	CTTGCACGAAG	ACGCTTGGCTGCGAAGACGCTTTGACCCGATACCGCTTTATTCGTCGCG	360			
Db	26397	CTTGCACGAAG	ACGCTTGGCTGCGAAGACGCTTTGACCCGATACCGCTTTATTCGTCGCG	26338			
QY	361	TTGGAACACGATG	TTTATGACACGTCCTGACCTCGCCCTCGGCGTGGCGATTACTGCGGG	420			
Db	26337	TTGGAACACGATG	TTTATGACACGTCCTGACCTCGCCCTCGGCGTGGCGATTACTGCGGG	26278			
QY	421	CGCGCTTTCGCGCT	GTGTGGAAGGGAACCTG	GGGGGACGCGCGGGCTATATATCTCCGGA	480		
Db	26277	CGCGCTTTCGCGCT	GTGTGGAAGGGAACCTG	GGGGGACGCGCGGGCTATATATCTCCGGA	26218		
QY	481	AAAGCGATG	CGGGTTTTCTTGACAGGTTG	CGCCCTCGCGCCGACGAAGGCTGCACCCC	540		
Db	26217	AAAGCGATG	CGGGTTTTCTTGACAGGTTG	CGCCCTCGCGCCGACGAAGGCTGCACCCC	26158		
QY	541	GTGCAATCTGATG	ATGTTTTCACGCAATTTTTTTCAGACAGGGAAGAAATG	CGCGTGTTCACAGCTTC	600		
Db	26157	GTGCAATCTGATG	ATGTTTTCACGCAATTTTTTTCAGACAGGGAAGAAATG	CGCGTGTTCACAGCTTC	26098		
QY	601	AATCCCGCTTTGTC	CGCCCAAGAGCTGCATTATGSCAAAGTTTTCAGACCAAAACAGCGCA	660			
Db	26097	AATCCCGCTTTGTC	CGCCCAAGAGCTGCATTATGSCAAAGTTTTCAGACCAAAACAGCGCA	26038			
QY	661	TTGGGCAAGCTGAT	TGAAACGACCGGCTTCTTGAACGCAACGAAAGCGCGCATTC	720			
Db	26037	TTGGGCAAGCTGAT	TGAAACGACCGGCTTCTTGAACGCAACGAAAGCGCGCATTC	25978			
QY	721	CCGCGCAACACAT	TTCAAACACCGGCTGATCGGCGCTTGACCAAAATTCAGCAGGGAAGG	780			
Db	25977	CCGCGCAACACAT	TTCAAACACCGGCTGATCGGCGCTTGACCAAAATTCAGCAGGGAAGG	25918			
QY	781	GAAGAAACGCGG	CAAGGCGCGAAGCAGTTGATGCTTTTCAATTA	828			

DB	25917	GAAAAAGCGCGGAAAGCGCGAAGCATTTGATGCTTTCCATATA	25870
RESULT 3			
AX044034/c			
LOCUS	AX044034	349980 bp	DNA linear PAT 24-NOV-2000
DEFINITION	Sequence 113 from Patent WO006791.		
ACCESSION	AX044034		
VERSION	AX044034.1	GI:11342918	
KEYWORDS			
SOURCE			
ORGANISM	Neisseria meningitidis		
	Neisseria meningitidis		
	Bacteriata; Proteobacteria; Betaproteobacteria; Neisseriales;		
	Neisseriaceae; Neisseria.		
REFERENCE			
AUTHORS	1		
	Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,		
	Maignani, V., Galeotti, C., Mora, M., Ratti, G., Scarcell, M.,		
	Scariato, V., Rappunli, R., Frazer, C. M. and Grandi, G.		
	Neisseria genomic sequences and methods of their use		
	Patent: WO 006791-A 113 09-NOV-2000;		
JOURNAL	CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)		
FEATURES			
source	1. 349980		
	/organism="Neisseria meningitidis"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:487"		
	/note="sequence too long, cut in 8 pieces."		
	seq 1: 1 to 34980 34980 bases		
	seq 108: 300001 to 649980 349980 bases		
	seq 109: 600001 to 949980 349980 bases		
	seq 110: 900001 to 1249980 349980 bases		
	seq 111: 1200001 to 1549980 349980 bases		
	seq 112: 1500001 to 1849980 349980 bases		
	seq 113: 1800001 to 2149980 349980 bases		
	seq 114: 2100001 to 2272325 172325 bases"		
ORIGIN			
Query Match	99.8%; Score 826.4; DB 6; Length 349980;		
Best Local Similarity	99.9%; Pred. No. 2,6e-186;		
Matches 827; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	ATGCAAAACCAAGTATACGTTAGCTTCCGCGGAGAACGAGGCGGACATTTCCGAT	60
DB	226656	ATGCAAAACCAAGTATACGTTAGCTTCCGCGGAGAACGAGGCGGACATTTCCGAT	226597
QY	61	ACCTTGCGGAGGACGCGCATCCCGTTTCAAGTTTTCACGACATGATGCCGTGAAAG	120
DB	226596	ACCTTGCGGAGGACGCGCATCCCGTTTCAAGTTTTCACGACATGATGCCGTGAAAG	226537
QY	121	CTGGAACAGGACATGGCGGACCTGCTCCCGGCTGTGGCGGACCCCTATTGAGCGGA	180
DB	226536	CTGGAACAGGACATGGCGGACCTGCTCCCGGCTGTGGCGGACCCCTATTGAGCGGA	226477
QY	181	GTCGAAAAAGCGCTGTTATGAGCAGCGCGGATTTGTAAGAGCAGCATTTGAGAAAGT	240
DB	226476	GTCGAAAAAGCGCTGTTATGAGCAGCGCGGATTTGTAAGAGCAGCATTTGAGAAAGT	226417
QY	241	CTGCGGTTATATCAACGATTTGAGAGCAGCTTTTCTCGCGGAAAGTGAGAAAAATTC	300
DB	226416	CTGCGGTTATATCAACGATTTGAGAGCAGCTTTTCTCGCGGAAAGTGAGAAAAATTC	226357
QY	301	CTTGCGGAAGAGCGCTTGCGTGCAGAAAGGCTTTGACCGGAGTACCGGCTTATGTCGCG	360
DB	226356	CTTGCGGAAGAGCGCTTGCGTGCAGAAAGGCTTTGACCGGAGTACCGGCTTATGTCGCG	226297
QY	361	TTGGAACAGATGTTATGACAGCTCTGACCTTCGCGCTGCGGAGTTACTGCGAGG	420
DB	226296	TTGGAACAGATGTTATGACAGCTCTGACCTTCGCGCTGCGGAGTTACTGCGAGG	226237
QY	421	CGGCGCTTTCCGCTGTGGAAGGCAACATCGGGGAGACGGGGGCTATATCATTTCCGA	480
DB	226236	CGGCGCTTTCCGCTGTGGAAGGCAACATCGGGGAGACGGGGGCTATATCATTTCCGA	226177

481 AAAGCGATGCGGTTTCTGACAGGTTTGCAGCCCTGCGCCGGAAGGCTGCACCCC 540
DB 226176 AAAGCGATGCGGTTTCTGACAGGTTTGCAGCCCTGCGCCGGAAGGCTGCACCCC 226117
QY 541 GTCGATCTGATGATGTTTCAGCCGATTTTTCGACAGGGAAGAAATGCCGTTTGCAGCTC 600
DB 226116 GTCGATCTGATGATGTTTCAGCCGATTTTTCGACAGGGAAGAAATGCCGTTTGCAGCTC 226057
QY 601 AATCCCGCTTGGCCCGCCAGAGCTGCATTAAGCCAGTTTTCAGACCCAAAACAGCCGA 660
DB 226056 AATCCCGCTTGGCCCGCCAGAGCTGCATTAAGCCAGTTTTCAGACCCAAAACAGCCGA 225997
QY 661 TTGGGAGCGCTGATTCGACACGACCGCTCTGTAACCGCAAAAGGCGCGATTC 720
DB 225996 TTGGGAGCGCTGATTCGACACGACCGCTCTGTAACCGCAAAAGGCGCGATTC 225937
QY 721 CCCGCCAACAATTCAAAACCGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 225936 CCCGCCAACAATTCAAAACCGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 225877
QY 781 GAAAAACGCGCGCAAGGCGCGCAACAGTTCAATGTCGCTTTCCAATTA 828
DB 225876 GAAAAACGCGCGCAAGGCGCGCAACAGTTCAATGTCGCTTTCCAATTA 225829

RESULT 4 4827 bp DNA linear BCT 02-JAN-2003
LOCUS AY039684
DEFINITION *Neisseria meningitidis* strain B305-Tr4 11pooligosaecharide
glycosyltransferase locus, complete sequence.
ACCESSION AY039684
VERSION AY039684.1 GI:21654774
KEYWORDS
SOURCE *Neisseria meningitidis*
ORGANISM *Neisseria meningitidis*
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 4827)
AUTHORS Zhu, P., Tsai, C.M. and Frasch, C.E.
TITLE Immunologic and genetic characterization of lipooligosaccharide
variants in a *Neisseria meningitidis* serogroup C strain
JOURNAL FEMS Immunol. Med. Microbiol. 34 (3), 193-200 (2002)
PUBMED 12423771
AUTHORS Zhu, P., Lacerda, M.D., Tsai, C.M. and Frasch, C.E.
TITLE Direct Substitution
JOURNAL Submitted (12-JUN-2001) LBP/DBPAP, CBER/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source Location/Qualifiers
1..4827
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="B305-Tr4"
/db_xref="taxon:487"
38..439
/note="OrfA; allele 2"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK85136.1"
/db_xref="GI:21654775"
/translation="MPSEAFRRHAYRENKLOPLSVLICAVNVEKYFAOSLAAYVNO
TWRNLEILVDDSGDTALAKDROKDSRKLIAHRLNLRKQORDSPANTYKHK
LIRALTKISREKRRQRREGLIKIIVFG"
451..1383
/gene="1gtc"
/gene="1gtc"
/allele="6"
451..1383
/gene="1gtc"
/allele="6"
/note="1gtc"

Query Match 97.7%; Score 808.8; DB 1; Length 4827;
Best Local Similarity 96.6%; Pred. No. 3.5e-182;
Matches 816; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
ORIGIN
1 ATGCAAAACGCGTATCAGCTTACGCTTCGCGCAGACGAGGCGCATTCGCCGAT 60
DB 2478 ATGCAAAACGCGTATCAGCTTACGCTTCGCGCAGACGAGGCGCATTCGCCGAT 2537
QY 61 ACCTTGGCAGGACGCGATTCGCTTTCAGTTTTCAGACGACGATGATGCGCTGAAAG 120
DB 2538 ACCTTGGCAGGACGCGATTCGCTTTCAGTTTTCAGACGACGATGATGCGCTGAAAG 2597

/codon_start=1
/transl_table=11
/product="alpha 1,4 galactosyltransferase"
/protein_id="AAK85137.1"
/db_xref="GI:21654776"
/translation="MDIVFAADDDVAAVLAIVAAKSVAAHPDTEIRPHVLDAIISBAN
RAAVANLRGGNIRFTIDVNPEDFAGPPLNIRHSITTYARLKIAGEYADDDKYLXLD
IDVVRDLSKRLMTDLDNVLGACIDIFERQAAKYKIKGMAAGEYFNAGVLLINL
KKRHDIKMAKCEVDEYKDVQYQDDIINGLPKGSVCYANRFPNPTNDLPMAN
RPAASHDPLRYDRTYTMAPVAVSHYCGAPMWRHDCDAGARFTBLAGSLTSVEE
WRGLAVHRVLPPTKRLQWRRLISARFLKLY"
1438..2436
/gene="1gtc"
/allele="9"
1438..2436
/gene="1gtc"
/allele="9"
/note="1gtc"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="beta 1,3 N-acetylglucosaminyltransferase"
/protein_id="AAK85138.1"
/db_xref="GI:21654777"
/translation="MQSIVSVLICAVNVEKYFAOSLAAYVNOQWRNLEILVDDGSTD
GTLIAKDFKDRSRIRKILAOASGLIPSLNIGLDELAKSGMEGYARTADADIAAP
DWIRKIVGEMEKRSITAMGAMLEVLSEKQGNRLAHHRRKIKKKPTTHREDIADFP
PFGNPHNNTIMRVSVDGLRYNTERDLAGSMDGYKTPRDSIEFROICAVAYELLKHL
HANOVSSKYSRQHEITMAOIKTRANDPLFDSMGKTPRDSIEFROICAVAYELLKHL
PDEPERRKRLFYCCFKRTDPLPAGAWLDFPADSRMRLLFTROYFGILHRLNKR"
2478..3305
/gene="1gtcB"
/allele="6"
2478..3305
/gene="1gtcB"
/allele="6"
/note="1gtcB"
/codon_start=1
/transl_table=11
/product="beta 1,4 galactosyltransferase"
/protein_id="AAK85139.1"
/db_xref="GI:21654778"
/translation="MOMNVISLASAARRAHADTPRRHGIPFPFALMPSELEQA
MAELVPGLSAPHYISGVKACFMSHAYLWQALDEGLPYIAVEFDVLLGEGAKFLA
EDAWLORFDPDSAPIVRLTFEMHVLITSGVADYCGARAPLLESHEWGTAGYIISR
KAMRFLDRFALPPEGHPVDLMFEDFDRBEGVPYCOLNPALCAQELHYAKHDON
SALGSLIEHRLNLRKQORDSPANTYKHLIRALTKISREKRRQRREGLIIVFG"
3454..4260
/note="OrfB1; allele 1"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK85140.1"
/db_xref="GI:21654779"
/translation="MOMNVISLASAARRAHADTPRRHGIPFPFALMPSELEQA
MAELVPGLSAPHYISGVKACFMSHAYLWQALDEGLPYIAVEFDVLLGEGAKFLA
EDAWLORFDPDSAPIVRLTFEMHVLITSGVADYCGARAPLLESHEWGTAGYIISR
KAMRFLDRFALPPEGHPVDLMFEDFDRBEGVPYCOLNPALCAQELHYAKHDON
SALGSLIEHRLNLRKQORDSPANTYKHLIRALTKISREKRRQRREGLIIVFG"
4260..4827
/note="OrfB1; allele 1"

QY	121	CTGGAAACAGGCAATGAGCGGAACCTGCTCCCCGGGCTGTGGGCGACACCCCTAATTGAGCGGA	180
Db	2598	CTGGAAACAGGCAATGAGCGGAACCTGCACTCCGGTGTGTGGCGCACCCCTAATTGAGCGGA	2657
QY	181	GTGGAAAAAGCCTGCTTTATGAGCCACGCGCTATTTGTGAAGCAGGCAATTGACAGAGT	240
Db	2658	GTGGAAAAAGCCTGCTTTATGAGCCACGCGCTATTTGTGAAGCAGGCAATTGACAGAGT	2717
QY	241	CTGCCGTATATCACCGTATTTGAGACGACGTTTTACTCGCGGAAGTGAAGAAAAATTC	300
Db	2718	CTGCCGTATATGCGCGTATTTGAGACGACGTTTTACTCGCGGAAGCGCGGAAGAAATTC	2777
QY	301	CTTGTGTGAAGAGCTTTGGCTGCAAGAACGCTTTGACCCGGAATCCGCTTTATGCTCCG	360
Db	2778	CTTGTGTGAAGAGCTTTGGCTGCAAGAACGCTTTGACCCCGAATCCGCTTTATGCTCCG	2837
QY	361	TTGAAAAAGATTTTATGACAGTCTCGACCTCGCGCTCGGCTGGGCGGATTAATCGCGGG	420
Db	2838	TTGAAAAAGATTTTATGACAGTCTCGACCTCGCGCTCGGCTGGGCGGATTAATCGCGGG	2897
QY	421	CGCGCTTTCCGCTGTGTAAGAAAGCAACACTGGGGGAGCGGCGGCTATATCATTTCCGGA	480
Db	2898	CGCGCTTTCCGCTGTGTAAGAAAGCAACACTGGGGGAGCGGCGGCTATATCATTTCCGGA	2957
QY	481	AAAGCGATGCGGTTTTTCTCTGACAGGTTTTGCGCGCTCGCGCCGGAAGGGCTGCACCC	540
Db	2958	AAAGCGATGCGGTTTTTCTCTGACAGGTTTTGCGCGCTCGCGCCGGAAGGGCTGCACCC	3017
QY	541	GTGCAATCGATGATNTTTCAGCGATTTTTTTCACAGGGAAGGAATGCCGGTTTGCAAGTC	600
Db	3018	GTGCAATCGATGATNTTTCAGCGATTTTTTTCACAGGGAAGGAATGCCGGTTTGCAAGTC	3077
QY	601	AATCCCGGCTTGTGGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGCAACAAACAGCGCA	660
Db	3078	AATCCCGGCTTGTGGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGCAACAAACAGCGCA	3137
QY	661	TTGGGACGCTTGATGAAACGACCGCGCTCTTGAAACGCAACAGCAAAAGCGCGATTC	720
Db	3138	TTGGGACGCTTGATGAAACGACCGCGCTCTTGAAACGCAACAGCAAAAGCGCGATTC	3197
QY	721	CCCGGCAACATTCAAACACCGCTGATCCGCGCTTGACCAAAATTAAGCAGGGAAGG	780
Db	3198	CCCGGCAACATTCAAACACCGCTGATCCGCGCTTGACCAAAATTAAGCAGGGAAGG	3257
QY	781	GA AAAAACGCGGCAAAAGCGGCAACAGTTCAATTGCGCTTCCAAATA	828
Db	3258	GA AAAAACGCGGCAAAAGCGGCAACAGTTCAATTGCGCTTCCAAATA	3305

RESULT 5	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	TITLE	JOURNAL	REFERENCE	FEATURES
AF470659	AF470659	Neisseria meningitidis strain M981 1gt-1 locus, complete sequence.	AF470659	AF470659.1	GI:21069184						
		3335 bp	DNA	linear	BCT 22-MAY-2002						
							Neisseria meningitidis				
							Neisseria meningitidis				
							Bacteria; Proteobacteria;				
							Betaproteobacteria; Neisseriales;				
							Neisseriaceae; Neisseria.				
							1 (bases 1 to 3335).				
							Zhu, P., Klutch, M. J., Bash, M. C., Tsang, R. S., Ng, L. K. and Tsai, C. M.				
							Genetic diversity of three 1gt loci for biosynthesis of				
							1-pooligossacharide (LOS) in Neisseria species				
							Microbiology 148 (Pt 6), 1833-1844 (2002)				
							2 (bases 1 to 3335)				
							Zhu, P.				
							Direct Submission				
							Submitted (22-JUN-2002)				
							DBPAP, CBER/FDA, 8800 Rockville Pike,				
							Bethesda, MD 20892, USA				
							Location/Qualifiers				

source	1.. 3335
gene	/organism="Neisseria meningitidis" /mol_type="genomic DNA" /strain="M981" /db_xref="taxon:487" 104..1192 /gene="lgtA" /allele="13" 104..1192 /gene="lgtA" /allele="13" /note="glycosyl transferase" /codon_start=1 /transl_table=11 /product="lgtA" /protein_id="AA03861.1" /db_xref="GI:21069185" /translation="MPSEAFPRRRAHYRENNKLOPLVSVLTCAYNEKYFAQSLAAVNC TWRNLDLIVDDGSDGTGLTAQRFQEDDGRIRILAOFRNSGLPSINIGLDELAKSGS MGEYATRDADDIAAPDWIKIVGEMMEDRSITMGAMLVLSRKQGNRLAHHHRC KIKKPPRHEDLADPPRGNIPNNMTSRSVIDGLRYTTERDMADRYDPWYDVSLE LGRLATYPRALVKTIRLANOVSSKYSITQHEIAGIQIARNDPLQSNQFKTRDVSLE YRQIKAVAYELERKLPEDEDERARRELTYOCFKRTDPAPAGAMLDFADGKRRLLFTY RQYFGLARLLIKNNRQARSQAKGEOR1"
CDS	1192..2019 /gene="lgtB" /allele="13" 1192..2019 /gene="lgtB" /allele="13" /note="glycosyl transferase" /codon_start=1 /transl_table=11 /product="lgtB" /protein_id="AA03862.1" /db_xref="GI:21069186" /translation="MOMHVISLAAERRAHITDPFGVRCIPPOFPDAMPSERLLEQ MAELVPGASHPYGVGEKACFMSHAYLAKMOALDEGLPYIAVFEDDVLIGSAEKPLA KAMLEOERPDSDSAFIVRLIETMFMHVLTPSGVADYCGRAPLIESHWGAYGIIIS KAMPPLDRPALPPEGLHPYDLMLPSPDPREGVPPYCOALPALCAQELIRAKKHDDQ SALCSLIEHDLRLNRKQOMRDSPPANTFRRLIRALTYKISREKRRQRRQEOIVPFO 2168..3010 /gene="lgtE" /allele="8" 2168..3010 /gene="lgtE" /allele="8" /note="glycosyl transferase" /codon_start=1 /transl_table=11 /product="lgtE" /protein_id="AA03863.1" /db_xref="GI:21069187" /translation="MOMHVISLAAERRAHITDPFGRHDIPOFPDAMPSERLNRM MAELVPGAKOHLSEVEKACFMSHAYLWFOALDESIPYIAVFEDDVLIGSAEQPLA EDTWLERPDSDSAFIVRLIETMFMHVLTPSGVADYCGRAPLIESHWGAYGIIIS KAMPPLDRPALPPEGLHPYDLMLPSPDPREGVPPYCOALPALCAQELIRAKKHDDQ SALCSLIEHDLRLNRKQOMRDSPPANTFRRLIRALTYKISREKRRQRRQEOIVPFO ISF"
gene	2168..3010 /gene="lgtE" /allele="8" 2168..3010 /gene="lgtE" /allele="8" /note="glycosyl transferase" /codon_start=1 /transl_table=11 /product="lgtE" /protein_id="AA03863.1" /db_xref="GI:21069187" /translation="MOMHVISLAAERRAHITDPFGRHDIPOFPDAMPSERLNRM MAELVPGAKOHLSEVEKACFMSHAYLWFOALDESIPYIAVFEDDVLIGSAEQPLA EDTWLERPDSDSAFIVRLIETMFMHVLTPSGVADYCGRAPLIESHWGAYGIIIS KAMPPLDRPALPPEGLHPYDLMLPSPDPREGVPPYCOALPALCAQELIRAKKHDDQ SALCSLIEHDLRLNRKQOMRDSPPANTFRRLIRALTYKISREKRRQRRQEOIVPFO ISF"
CDS	2168..3010 /gene="lgtE" /allele="8" 2168..3010 /gene="lgtE" /allele="8" /note="glycosyl transferase" /codon_start=1 /transl_table=11 /product="lgtE" /protein_id="AA03863.1" /db_xref="GI:21069187" /translation="MOMHVISLAAERRAHITDPFGRHDIPOFPDAMPSERLNRM MAELVPGAKOHLSEVEKACFMSHAYLWFOALDESIPYIAVFEDDVLIGSAEQPLA EDTWLERPDSDSAFIVRLIETMFMHVLTPSGVADYCGRAPLIESHWGAYGIIIS KAMPPLDRPALPPEGLHPYDLMLPSPDPREGVPPYCOALPALCAQELIRAKKHDDQ SALCSLIEHDLRLNRKQOMRDSPPANTFRRLIRALTYKISREKRRQRRQEOIVPFO ISF"
ORIGIN	Query Match 96.7%; Score 800.8; DB 1; Length 3335; Best Local Similarity 97.9%; Pred. No. 2.8e-180; Matches 811; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY	1 ATGCAAAACCGATTATCAGTTAGCTTCGCGCGCAGAAAGCGAGGCGCATTTGCCGAT 60 DB 1192 ATGCAAAACCGATTATCAGTTAGCTTCGCGCGCAGAAAGCGAGGCGCATTTGCCGAT 1251 QY 61 ACCTTCGGCAGGCGGCGATCCGCTTCAGTTTTCGACCGCATGATCCGCTGGAAG 120 DB 1252 ACCTTCGGCAGGCGGCGATCCGCTTCAGTTTTCGACCGCATGATCCGCTGGAAG 1311

QY	121	CTGGAACAGGCATATGGCGAAGCTGTCCTCCCGCTTGTGCGCGACCCCTATTGAGCGGA	180		
Db	1312	CTGGAACAGGCATATGGCGAAGCTGTCCTCCCGCTTGTGCGCGACCCCTATTGAGCGGA	1371		
QY	181	GTGGAAGAAAGCCTGCTTTATGAGCCACGCGATATTGGAAGAGGCAATTGACGAAGGT	240		
Db	1372	GTGGAAGAAAGCCTGCTTTATGAGCCACGCGATATTGGAAGAAAGGCAATTGACGAAGGT	1431		
QY	241	CTGCGCTATATTCACCGTATTTGAGAGCAGAGTTTACTGCGGCAAGTAGAGAAAATTTC	300		
Db	1432	CTGCGCTATATTCACCGTATTTGAGAGCAGAGTTTACTGCGGCAAGTAGAGAAAATTTC	1491		
QY	301	CTTGCGAAGACGCTTGCTGCAAGACGCTTGAACCCGAGTACCCCTTTATGTCGCCG	360		
Db	1492	CTTGCGCAAGACGCTTGCTGCAAGACGCTTGAACCCGAGTATCCGCTTTATGTCGCC	1551		
QY	361	TTGGAAGACGATGTTTATGCAAGTCTGACCTCGCCCTCGGCGTGCAGATTACTGCGGG	420		
Db	1552	TTGGAAGACGATGTTTATGCAAGTCTGACCTCGCCCTCGGCGTGCAGATTACTGCGGG	1611		
QY	421	CGCGCCTTCCGCTGTTGGAAGACGACACTGGGGGACGCGCGGCTATATCATTTTCCGA	480		
Db	1612	CGCGCCTTTCGCTGTGTGGAAGACGACCTGGGGGACGCGCGGCTATATCATTTTCCGA	1671		
QY	481	AAAGCATGCGGCTTTTCTTGAACAGGTTTGCCGCTCGCCCGACGAGGCTGCACCCC	540		
Db	1672	AAAGGGAAGCGGTTTTTCTTGACAGGTTTGCCGCTCGCCCGACGAGGCTGCACCCC	1731		
QY	541	GTCGATCTGATGATTTGACGCGATTTTTCGACAGGGAAGAAATGCGGTTTCCGACGCTC	600		
Db	1732	GTCGATCTGATGATTTGACGCGATTTTTCGACAGGGAAGAAATGCGGTTTGCGACGCTC	1791		
QY	601	AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAAGTTTACAGCAACAAACAGCGCA	660		
Db	1792	AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAAGTTTACAGCAACAAACAGCGCA	1851		
QY	661	TTGGGAGCGCTGATTCGACAGCAGCGCTCTGGAACCGCAACAGCAAGGCGCATTC	720		
Db	1852	TTGGGAGCGCTGATTCGACAGCAGCGCTCTGGAACCGCAACAGCAATGGCGCATTC	1911		
QY	721	CCGCGCAACACATTCGACCAACCGCTGATTCGCGCGCTTGAACCAAAATCAGAGGGAAG	780		
Db	1912	CCCGCAACACATTCGACCAACCGCTGATTCGCGCGCTTGAACCAAAATCAGAGGGAAG	1971		
QY	781	GAAAAAGCGCGCAAGGCGGCAACGTTGATTTGCTTTCCATTA	828		
Db	1972	GAAAAAGCGCGCAAGGCGGCAACGTTGATTTGCTTTCCATTA	2019		
RESULT 6	AF240672	2959 bp	DNA	linear	BCT 01-FEB-2001
LOCUS	AF240672				
DEFINITION	Neisseria subflava N-acetylglucosamine transferase (lgtA), galactosyl transferase (lgtB), and galactosyl transferase (lgtB) genes, complete cds.				
ACCESSION	AF240672				
VERSION	AF240672.1 GI:9963919				
KEYWORDS					
SOURCE	Neisseria subflava				
ORGANISM	Neisseria subflava				
REFERENCE	Neisseria subflava; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.				
AUTHORS	1 (bases 1 to 2959)				
TITLE	Arking, D., Tong, Y., and Stein, D. C.				
JOURNAL	Analysis of lipooligosaccharide biosynthesis in the Neisseriaceae				
PUBMED	J. Bacteriol. 183 (3), 934-941 (2001)				
REFERENCE	2 (bases 1 to 2959)				
AUTHORS	Stein, D.C. and Tong, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAR-2000) Cell Biology and Molecular Genetics,				
FEATURES	University of Maryland, College Park, MD 20742, USA				
	Location/Qualifiers				

	source	1..2959	/organism="Neisseria subflava"
	gene	/mol_type="genomic DNA"	
	CDS	/db_xref="taxon:28449"	
		19..1065	/gene="lgtA"
		19..1065	/gene="lgtB"
		19..1065	/note="lgtA"
		/note="lgtB"	
		/codon_start=1	
		/transl_table=11	
		/product="N-acetylglucosamine transferase"	
		/protein_id="AAG09764.1"	
		/db_xref="GI:9963920"	
		/translation="MPSEAFRRHRAVRENKLOPLVSVLICAYNVEKYFAGSLAAYVQ	
		TWRLDILIVDDGSDTGLIARRFOBQGRIRILAOFRNSGLIPSLNIGLDELAKSG	
		MGSEYIARDADDIAPDWIERKIVGEMKDRSIIAMGAMLEYSKDNDRHHRHG	
		KIMKPTREDIADFPFGNPVHNNTMRSVLDGGLRYNTERBDMADYQFWYDVK	
		LGRLAYVEALVKYRIHANOVSSTYSVROHIEIAGIOKTARNDFLOMGFPTPDSLE	
		RYQIKAVYELLEGHLPEBDFERRARPLVQCPRKTDITLPAQAMLDPAADGRMLFTL	
		ROYRILHRLKNR"	
	gene	1107..1934	/gene="lgtB"
	CDS	1107..1934	/gene="lgtB"
		/note="lgtB"	
		/codon_start=1	
		/transl_table=11	
		/product="galactosyl transferase"	
		/protein_id="AAG09765.1"	
		/db_xref="GI:9963921"	
		/translation="MONHVISLAASBRRAHITDPFGVRGIPQFPDMLMPSRELEOV	
		MAELVPGISAPVLSGVKACFMGSAVLAKQALDEGLPYIVFEDDVLGGAGKPLA	
		EDAWLOERPDSDAFIVRIETMFWHTSPSGVADYGRALPILSEHMGATITISK	
		KAIRPFLDRPALPPEGLHPYDLMFSDPFRBGVPYQOLNPAICAGELHAKHDON	
		SALGSLIBHDLRLNRKQMDSPANTFRRLIRALTYKISRERKRRQREQIVPFO"	
		2083..2925	/gene="lgtB"
	gene	2083..2925	/gene="lgtB"
	CDS	2083..2925	/gene="lgtB"
		/note="lgtB"	
		/codon_start=1	
		/transl_table=11	
		/product="galactosyl transferase"	
		/protein_id="AAG09766.1"	
		/db_xref="GI:9963922"	
		/translation="MONHVISLAASBRRAHITDPFGVRGIPQFPDMLMPSRELEOV	
		MAELVPGIAXOHLSEVKEACFMGSAVLAKQALDEGLPYIVFEDDVLGGAGKPLA	
		EDTWLERPDSDAFIVRIETMFWHTSPSGVADYGRALPILSEHMGATITISK	
		BAKPFLERFALPAEWIKAYDMWMTFFPDKEGMPVYQVNPALCTGELHAKFLSKN	
		SMGSDLEKREORRRRRSLIKVWFILKRALGKPRBKRMERORQAELEKAYGRV	
		ISFK"	
ORIGIN			
Query Match	95.6%;	Score 791.2;	DB 1; Length 2959;
Best Local Similarity	97.2%;	Pred. No. 5.5e-178;	
Matches	805;	Conservative 0;	Mismatches 23; Indels 0; Gaps 0;
QY	1	ATGCAAAACGAGTATCAGCTTAGCTTCGCGCCGACGACGAGCGGACGATTCGCGAT	60
Db	1107	ATGCAAAACGAGTATCAGCTTAGCTTCGCGCCGACGACGAGCGGACGATTCGCGAT	1166
QY	61	ACCTTGGGAGGACGAGCATCCGCTTTCAGATTTTTCAGACGACGATGCGCTGGAAGG	120
Db	1167	ACCTTGGGAGGACGAGCATCCGCTTTCAGATTTTTCAGACGACGATGCGCTGGAAGG	1226
QY	121	CTGGAACAGGCAAGGCGGAACCTGTCGCGGCTTGTGCGGACGCCCTATTGAGCGGA	180
Db	1227	CTGGAACAGGCAAGGCGGAACCTGTCGCGGCTTGTGCGGACGCCCTATTGAGCGGA	1286
QY	181	GTGGAAGAAAGCCTGCTTTATGAGCCACGCGATATTGGAAGAGGCAATTGACGAAGGT	240
Db	1287	GTGGAAGAAAGCCTGCTTTATGAGCCACGCGATATTGGAAGAGGCAATTGACGAAGGT	1346

QY	241	CGCGCGATATATACCGGATNTTGAAGACGAGTTTAACTCCGGAGAAAGTGAGAAAATTC	300	
Db	1466	CTGCCGTATATGCGCGATATTTAAGATATGTCTTGCTGGGAAAGGCGCGAATAATTC	1525	
QY	301	CTTGCAGAAAGCGCTTGGCTGCGAAGACGCTTTGACCCGCGAATACCGCTTATCGTCCG	360	
Db	1526	CTTGCAGAAAGCGCTTGGCTGCGAAGACGCTTTGACCCGCGAATACCGCTTATCGTCCG	1585	
QY	361	TTGGAAACGATNTTATATGACGTCTTGACCTTGCCCTCCGGCGTGTGGCGAATTACTGGCG	420	
Db	1586	TTGGAAACGATNTTATATGACGTCTTGACCTTGCCCTCCGGCGTGTGGCGAATTACTGGCG	1645	
QY	421	CGCGCTTTCCGCTGTWGAAGCAACATCGAGGAGACGCGCGGCTATATCATTTCCCGA	480	
Db	1646	TGCGCTTCCGCTGTWGAAGCAACATCGAGGAGACGCGCGGCTATATCATTTCCCGA	1705	
QY	481	AAAGCGATGCGGTTTTTTCTTGACAGGTTTTGCGCCCTGCGCCGAAAGGCTGACCCC	540	
Db	1706	AAAGCGATGCGGTTTTTTCTTGACAGGTTTTGCGCCCTGCGCCGAAAGGCTGACCCC	1765	
QY	541	GTGCACTGATGATGTTCAGGCAATTTTTTCGACAGGAAAGAAATCGGTTTGCGAGTC	600	
Db	1766	GTGCACTGATGATGTTCAGGCAATTTTTTCGACAGGAAAGAAATCGGTTTGCGAGTC	1825	
QY	601	AATCCGCTTGTGTGCGCCCAAGAGCTGCATTAATGCCAATTTTCAGACCAAAACGCGCA	660	
Db	1826	GATCCCGCTTGTGTGCGCCCAAGAGCTGCATTAATGCCAATTTTCAGACCAAAACGCGCA	1885	
QY	661	TTGGGACAGCTGATGGAACAGACGCGCTCCCTGGAACCGCAACAGCAAGCGCGATTC	720	
Db	1886	TTGGGACAGCTGATGGAACAGACGCGCTCCCTGGAACCGCAACAGCAATGCGCGATTC	1945	
QY	721	CCCGCAACACATTCAAACACCGCTGATCCGCGCTTGACCAAAATCAGACGAGGAAAG	780	
Db	1946	CCCGCAACACATTCAAACACCGCTGATCCGCGCTTGACCAAAATCAGACGAGGAAAG	2005	
QY	781	GAAGAAAGCGCGCAAGCGCGACAGTTCAATGTGCTTTCAATA	828	
Db	2006	GAAGAAAGCGCGCAAGCGCGACAGTTCAATGTGCTTTCAATA	2053	
RESULT 8	AF470658	3367 bp	DNA linear	BCT 22-MAY-2002
LOCUS	AF470658			
DEFINITION	Neisseria meningitidis strain 891 1gt-1 locus, complete sequence.			
ACCESSION	AF470658			
VERSION	AF470658.1	GI:21069180		
KEYWORDS				
SOURCE				
ORGANISM	Neisseria meningitidis			
	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;			
	Neisseriaceae; Neisseria.			
REFERENCE	1 (bases 1 to 3367)			
AUTHORS	Zhu, P., Kutuch, M.J., Bash, M.C., Tsang, R.S., Ng, L.K. and Tsai, C.M.			
TITLE	Genetic diversity of three 1gt loci for biosynthesis of			
	lipooligosaccharide (LOS) in Neisseria species			
JOURNAL	Microbiology 148 (Pt 6), 1833-1844 (2002)			
PUBMED	12055303			
REFERENCE	2 (bases 1 to 3367)			
AUTHORS	Zhu, P.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-JAN-2002) DBPAP, CBER/FDA, 8800 Rockville Pike,			
	Bethesda, MD 20892, USA			
FEATURES				
Source	location/Qualifiers			
	1..3367			
	/organism="Neisseria meningitidis"			
	/mol_type="genomic DNA"			
	/strain="891"			
	/db_xref="taxon:487"			
	136..1182			
	/gene="lgtA"			
	/allele="9"			
	136..1182			
CDS				

Query Match	95.0%	Score 786.4	DB 1	Length 3367
Best Local Similarity	96.9%	Pred. No. 7.7e-177		
Matches 802	Conservative 0	Mismatches 26	Indels 0	Gaps 0
ORIGIN				
QY	1	ATGCAAAACCAAGCTTATGAGCACTTCCCGCCGCAAGAACGAGGCGCAATATGCCCAT	60	
Db	1224	ATGCAAAACCAAGCTTATGAGCACTTCCCGCCGCAAGAACGAGGCGCAATATGCCCAT	1283	
QY	61	ACCTTCGCGAGGCAAGGCAATCCCGTTTCAGTTTTCGAGCACTGATGCCGCTTGAAGAAG	120	
Db	1284	ACCTTCGCGCGGCAAGGCAATCCCGTTTCAGTTTTCGAGCACTGATGCCGCTTGAAGAAG	1343	
QY	121	CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTTCGCGGCAACCCCTATTTGACGGA	180	
Db	1344	CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTTCGCGGCAATCCCTATTTGACGGA	1403	
QY	181	GTGGAAAAAGCTGCTTTATGAGCAAGCGGTATTTGGAAGCAAGGTTGGAAGAAAGT	240	
Db	1404	GTGGAAAAAGCTGCTTTATGAGCAAGCGGTATTTGGAAGCAAGGTTTGAAGAAAGT	1463	
QY	241	CTGCCGATATCACCGTATTTGAGAGCAAGCTTTTATCTCGCGAAGGTGAGGAAAAATTC	300	

Db 1464 CTGCGGATATGCGCGGATTTGGAAGATGATGCTCTGCGCGAAGCGCGGAGAAATTC 1523
Qy 301 CTGGCGAAGACGCTTGCGCAAGACGCTTGAACCCGGAATACCGGCTTATCGTCCG 360
Db 1524 CTGGCGAAGACGCTTGCGCAAGACGCTTGAACCCGGAATACCGGCTTATCGTCCG 1583
Qy 361 TTGGAACGATGTTATGACAGTCTGACCTCGCCCTCGCGCGTGGCGAATTAATCTGCGG 420
Db 1584 TTGGAACGATGTTATGACAGTCTGACCTCGCCCTCGCGCGTGGCGAATTAATCTGCGG 1643
Qy 421 CGGCGCTTCCGCTGTTGGAAGCGAACACTGGGCGAGCGCGGCTTATCATTTCCGA 480
Db 1644 CGGCGCTTCCGCTGTTGGAAGCGAACACTGGGCGAGCGCGGCTTATCATTTCCGA 1703
Qy 481 AAGCGATGCGGTTTTCCTGCAAGGTTTCCGCGCCCTCGCGCGCGGCGTGCACCC 540
Db 1704 AAGCGATGCGGTTTTCCTGCAAGGTTTCCGCGCCCTCGCGCGGCGTGCACCC 1763
Qy 541 GTGATCTGATGATGTTTTCGACGATTTTTCGACAGGGAAGAAATGCGGTTTTCGACGCTC 600
Db 1764 GTGATCTGATGATGTTTTCGACGATTTTTCGACAGGGAAGAAATGCGGTTTTCGACGCTC 1823
Qy 601 AATCCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTCAAGCAAGGCA 660
Db 1824 GATCCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTTCGACAGGCA 1883
Qy 661 TTGGGCGAGCTGATCCGAACGACCGCTCTCTGTAACCGCAAGCAAGGCGGATTC 720
Db 1884 TTGGGCGAGCTGATCCGAACGACCGCTCTCTGTAACCGCAAGCAAGGCGGATTC 1943
Qy 721 CCGCGCAACATTCGAACGCGCTGATCCGCGCTCTGTAACCGCAAGCAAGGCAAGG 780
Db 1944 CCGCGCAACATTCGAACGCGCTGATCCGCGCTCTGTAACCGCAAGCAAGGCAAGG 2003
Qy 781 GAAAAACGCGCGCAAGGCGCAAGCAATTCATTTGCTTCCATTA 828
Db 2004 GAAAAACGCGCGCAAGGCGCAAGCAATTCATTTGCTTCCATTA 2051

RESULT 9
AY134877 5216 bp DNA linear BCT 02-AUG-2003
LOCUS Neisseria subflava strain 85071 lgc-1 locus, complete sequence.
DEFINITION AY134877
ACCESSION AY134877
VERSION AY134877.1 GI:33413660
KEYWORDS
SOURCE
ORGANISM
Neisseria subflava
Neisseria subflava
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 5216)
Zhu, P. and Tsai, C.M.
Galactosyltransferase gene family in Neisseria: Implication for
understanding the evolution of prokaryotic enzymes responsible for
glycosylation
Unpublished
2 (bases 1 to 5216)
Zhu, P.
Direct Submission
Submitted (22-JUL-2002) LBP/DBPAP/CBER, FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source
1..5216
/organism="Neisseria subflava"
/mol_type="genomic DNA"
/db_xref="taxon:28449"
/note="strain 85071"
181..1227
/gene="lgcA"
/allele="21"
181..1227
/gene="lgcA"

CDS

/allele="21"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="LgtA"
/protein_id="AA08514.1"
/db_xref="GI:33413661"
/translation="MPSBAPRRHRAVREKLOPIVSLICAVNKKYPAQSLAAVYNO
TWRNLEILIVDGSSTDTGLAIKNDPOKRSRIKIILAOQNSGLIPSLNIGLDELVYKQ
GSEYARTADIDIAFPDWTEKIVGEMKORSIIANGAMLSVSEKQNRNLAHHHGG
KIKWKEPTQHEIDIAAVFPNGPIHNNMIMRSVTDGALNTERDAEDYQFWYDVSK
LRLAAYPPALVYRLHANQVSKSVRQHEIAQGIQKTAANDPLOSMPFKTEFDSLE
YKIKAVAYELLEKHLPEEDFERARFLYQCFKRTDTPPAGAWLDPADQMRRLFTL
ROYFGIRLRLLNKR"
1269..2108
/gene="lgcB"
/allele="18"
1269..2108
/gene="lgcB"
/allele="18"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="LgtB"
/protein_id="AA08515.1"
/db_xref="GI:33413662"
/translation="MONHVISLASAERRRAHIAUTFGSRGICIPQFPDALMPSEBLEQA
MELVPGLSAHPYLSGVEKACPMASHAVLMKQALDGLPIAVFEDVLLBEGKEKELA
BDAMLOERPDPSAFIVRLTFMFMVLSPSGVADYCGRAPFLBSEHMGTAAYIISR
KAMRPFDRFPALPPEGHLPVDMMLSPDFDEGMPVCOINPALCQELHAYKPHQON
SALGSLIHDRLLNKKQRDRSPATPFKRLIRALTKISREKRRQRREQLIKIIV
PQO"
2120..3057
/gene="lgcC"
/allele="8"
/note="glycosyl transferase"
/pseudo
3117..4135
/gene="lgcD"
/allele="6"
/note="glycosyl transferase"
/pseudo
4154..4960
/gene="lgcH"
/allele="13"
4154..4960
/gene="lgcH"
/allele="13"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="LgtH"
/protein_id="AA08516.1"
/db_xref="GI:33413663"
/translation="MONHVISLASAKERRRAHIAUTFGRHGIPFOPFDALMPSEBLEQA
MAELVPGISAHPIYLSGVEKACPMASHAVLMKQALDGLPIAVFEDVLLBEGKEKELA
EDTWLOERPDPSAFIVRLTFMFMVLSPSGVADYCGRAPFLBSEHMGTAAYIISR
KAMRPFDRFPALPPEGHLPVDMMLSPDFDEGMPVCOINPALCQELHAYKPHQON
ROESYRNRRSPVILKRALGKIGREIBARERKROKLEKHLGRHVPEP"

ORIGIN

Query Match 94.4%; Score 782; DB 1; Length 5216;
Best Local Similarity 97.5%; Pred. No. 8.8e-176;
Matches 794; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCGATTATGAGTTAGCTCCGCGGAGAACGAGGCGGACATTCGCAT 60
Db 1269 ATGCAAAACCGATTATGAGTTAGCTCCGCGGAGAACGAGGCGGACATTCGCAT 1328
Qy 61 ACCTTGGGAGGACGCGATCCGTTTTCGACGACATGATGCGGCTGAAAAG 120
Db 1329 ACCTTGGGAGGACGCGATCCGTTTTCGACGACATGATGCGGCTGAAAAG 1388

```

QY 121 CTGGAACAGGCAATGGCGGAACCTGTCCTCCGCGCTTGTCCGCGCAACCCCTATTGAGCGGA 180
Db 1389 CTGGAACAGGCAATGGCGGAACCTGTCCTCCGCGCTTGTCCGCGCAACCCCTATTGAGCGGA 1448
QY 181 GTGGAAGAAAGCCCTGTTTATGAGCCAGCGCGATTTGTGAGAGGCAATTGAGCAAGGT 240
Db 1449 GTGGAAGAAAGCCCTGTTTATGAGCCAGCGCGATTTGTGAGAGGCAATTGAGCAAGGT 1508
QY 241 CTGCGCTATATACCGTATTGAGAGCAAGCTTTTACTCGCGCAGAGGTAGAGAAATTC 300
Db 1509 CTGCGCTATATACCGTATTGAGAGCAAGCTTTTACTCGCGCAGAGGTAGAGAAATTC 1568
QY 301 CTGCGCAAGAGACGCTTGCTGCAAGACGCTTTGACCCCGGATPACCGCTTTATCGATCCG 360
Db 1569 CTGCGCAAGAGACGCTTGCTGCAAGACGCTTTGACCCCGGATPACCGCTTTATCGATCCG 1628
QY 361 TTGGAACAGATGTTTATGCAAGCTCTGACCTCGCCCTCGCGCGTGGCGGATTAATGCGGG 420
Db 1629 TTGGAACAGATGTTTATGCAAGCTCTGACCTCGCCCTCGCGCGTGGCGGATTAATGCGGG 1688
QY 421 CGGCGCTTCCGCTGTTGGAAGCGCAACTGGGGGACGCGCGCTATATCATTTCCCA 480
Db 1689 CGGCGCTTCCGCTGTTGGAAGCGCAACTGGGGGACGCGCGCTATATCATTTCCCA 1748
QY 481 AAAGCATGCGGCTTTTCTGCAAGAGTTTGCCTGCGCCGCGCAAGGCGCTGCAACCC 540
Db 1749 AAAGCATGCGGCTTTTCTGCAAGAGTTTGCCTGCGCCGCGCAAGGCGCTGCAACCC 1808
QY 541 GTGCAATCATATATGTTTCAAGCGATTTTTCAGACGGAAGAAATGCGGTTGGCCAGCTC 600
Db 1809 GTGCAATCATATATGTTTCAAGCGATTTTTCAGACGGAAGAAATGCGGTTGGCCAGCTC 1868
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCCAGTTTTCAGACCAACCAACAGCGCA 660
Db 1869 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCCAGTTTTCAGACCAACCAACAGCGCA 1928
QY 661 TTGGGAGAGGCTATGCAACAGACCGCTCTCTGAAACCGCAACAGCAAGGCGCATTTCC 720
Db 1929 TTGGGAGAGGCTATGCAACAGACCGCTCTCTGAAACCGCAACAGCAAGGCGCGATTTCC 1988
QY 721 CCCGCAACACATTCGAAACAGCGCGCTGATCCGCGCTTGACCAAAATCGACGAGGAAAG 780
Db 1989 CCCGCAACACATTCGAAACAGCGCGCTGATCCGCGCTTGACCAAAATCGACGAGGAAAG 2048
QY 781 GAAAAACGCGCGCAAGGCGCGCAACGTTCAATTG 814
Db 2049 GAAAAACGCGCGCAAGGCGCGCAACGTTCAATCG 2082

RESULT 10
AP470663 3246 bp DNA linear BCT 22-MAY-2002
DEFINITION Neisseria meningitidis strain 7880 1gt-1 locus, complete sequence.
ACCESSION AP470663
VERSION AP470663.1 GI:21069199
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 3246)
Zhu, P., Klutch, M.J., Bash, M.C., Tsang, R.S., Ng, L.K. and Tsai, C.M.
Genetic diversity of three 1gt loci for biosynthesis of
lipooligosaccharide (LOS) in Neisseria species
Microbiology 148 (Pt 6), 1833-1844 (2002)
JOURNAL
PUBMED
12055303
REFERENCE
2 (bases 1 to 3246)
Zhu, P.
Direct Submission
Submitted (22-JAN-2002) DBPAP, CBER/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
JOURNAL
FEATURES
Location/Qualifiers
1..3246

```

```

/organism="Neisseria meningitidis"
/mol_type="Genomic DNA"
/strain="7880"
/db_xref="taxon:487"
140..1184
/gene="1gtA"
/allele="7"
/pseudo
140..1184
/gene="1gtA"
/allele="7"
/note="1gtA; glycosyl transferase"
pseudo
/codon_start=1
/transl_table=11
1226..2065
/gene="1gtB"
/allele="7"
1226..2065
/gene="1gtB"
/allele="7"
/note="1gtB"
CDS
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/db_xref="GI:21069200"
/protein_id="AAM33872.1"
/db_xref="gi:21069201"
/translation="MONHVISLAAABRAHIAATFGARCIPOFPDAMPSEBLEQ
MAELVPGSAHPYLSGVEKCFMSHVAIVMKALDEGAPYIAVFEDDVILGAKRFLA
EDAMLOERFDOSAFIVLETFMHWLTPSGVADYGRAPLLESEHFGAGYISR
KAMRFIDRPAVLPSERLKAIDMPLFSSFLDKGATTCOLTPALCVGSSTLPOLKNG
RQESYRNRSFKVLLKRALGKIGRIERARERRKOKLERLGRHVVEFE"
PFO"
2214..3020
/gene="1gtH"
/allele="8"
2214..3020
/gene="1gtH"
/allele="8"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtH"
/protein_id="AAM33873.1"
/db_xref="GI:21069201"
/translation="MONHVISLAAABRAHIAADTFGRHIDIPQFPDAMPSEBLEQ
MAELVPGSAHPYLSGVEKCFMSHVAIVMKALDEGAPYIAVFEDDVILGAKRFLA
EDAMLOERFDOSAFIVLETFMHWLTPSGVADYGRAPLLESEHFGAGYISR
KAMRFIDRPAVLPSERLKAIDMPLFSSFLDKGATTCOLTPALCVGSSTLPOLKNG
RQESYRNRSFKVLLKRALGKIGRIERARERRKOKLERLGRHVVEFE"
ORIGIN
Query Match 93.7%; Score 775.6; DB 1; Length 3246;
Best Local Similarity 97.1%; Pred. No. 2.9e-174;
Matches 790; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 ATGCAAAACACGTTATAGCTTACGCTTCGCGCGCAAGCAAGCGCGCACATTCGCCGAT 60
Db 1226 ATGCAAAACACGTTATAGCTTACGCTTCGCGCGCAAGCAAGCGCGCACATTCGCCGCA 1285
QY 61 ACCTTCGCGAGGACGCGCATCCGTTTCAAGTTTTCAGACGACATGATCCGCTGAAAG 120
Db 1286 ACCTTCGCGCGCGCGCGATCCGTTTCAAGTTTTCAGACGACATGATCCGCTGAAAG 1345
QY 121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCTTGTGCGCGCAACCCCTATTGAGCGGA 180
Db 1346 CTGGAACAGGCAATGCGGAACTGTCCTCCGCTTGTGCGCGCAACCCCTATTGAGCGGA 1405
QY 181 GTGGAAGAAAGCCCTGTTTATGAGCAAGCGGCTATTTGGAAGAGGCAATTGGAAGAGT 240
Db 1406 GTGGAAGAAAGCCCTGTTTATGAGCAAGCGGCTATTTGGAAGAGGCAATTGGAAGAGT 1465
QY 241 CTGCGTATATACCGTATTGAGAGCAAGCTTTTACTCGCGCAAGGTGAGAAATTC 300

```

Db 1466 CTCCGATATCGCCGATTTGAGAGCAAGTTTACTCGCGCAAGCGCGAGAAATTC 1525
Qy 301 CTGGCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCCGATACCGCTTATCGTCGC 360
Db 1526 CTGGCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCCGATACCGCTTATCGTCGC 1585
Qy 361 TTGGAACGATGTTTATGACAGCTCTGACCTCGCCCTCGCGCGTGGCGGATTACTGCGG 420
Db 1586 TTGGAACGATGTTTATGACAGCTCTGACCTCGCCCTCGCGCGTGGCGGATTACTGCGG 1645
Qy 421 CGGCGCTTCCGCTGTTGGAAGCGAACACTGGGGGACGCGCGGCTATATCATTTCCCGA 480
Db 1646 CGGCGCTTCCGCTGTTGGAAGCGAACACTGGGGGACGCGCGGCTATATCATTTCCCGA 1705
Qy 481 AAGGCAATCGGCTTTTCTCGACAGGTTTCCGCGCCCTCGCGCGGCGGCTGACCCC 540
Db 1706 AAGGCAATCGGCTTTTCTCGACAGGTTTCCGCGCCCTCGCGCGGCGGCTGACCCC 1765
Qy 541 GTGCACTGATGATGTTTCAAGCGATTTTTCGACAGGGAAGGAATGCGGTTGCCAGCTC 600
Db 1766 GTGCACTGATGATGTTTCAAGCGATTTTTCGACAGGGAAGGAATGCGGTTGCCAGCTC 1825
Qy 601 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTATGCCAAGTTTCAAGCAAGCAAGCA 660
Db 1826 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTATGCCAAGTTTCAAGCAAGCAAGCA 1885
Qy 661 TTGGGAGAGCTGATTCGAACAGCAGCGCTCTCTGTAACCGCAAGCAAGCAAGCGGATTC 720
Db 1886 TTGGGAGAGCTGATTCGAACAGCAGCGCTCTCTGTAACCGCAAGCAAGCAAGCGGATTC 1945
Qy 721 CCCGCAACACATTTCAACAGCAGCTGATCCGCGCTTGAACCAATTCAGCGGGAAGG 780
Db 1946 CCCGCAACACATTTCAACAGCAGCTGATCCGCGCTTGAACCAATTCAGCGGGAAGG 2005
Qy 781 GAAAAAGCGCGCAAGGCGCGCAACAGTTCAATTG 814
Db 2006 GAAAAAGCGCGCAAGGCGCGCAACAGTTAATCG 2039

RESULT 11
AP470662
LOCUS AP470662 3289 bp DNA linear BCT 22-MAY-2002
DEFINITION Neisseria meningitidis strain M120 1gt-1 locus, complete sequence.
ACCESSION AP470662
VERSION AP470662.1 GI:21069195
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 3289)
Zhu,P., Klutchev,M.J., Baah,M.C., Tsang,R.S., Ng,L.K. and Tsai,C.M.
Genetic diversity of three 1gt loci for biosynthesis of
Microbiogeochemical 148 (Pt 6), 1833-1844 (2002)

JOURNAL
PUBMED
2 (bases 1 to 3289)
Zhu,P.
Direct Submission
Submitted (22-JAN-2002) DBPAP, CBBR/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source
1..3289
location/Qualifiers
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="M120"
/db_xref="taxon:487"
106..1152
/gene="1gtA"
106..1152
/allele="6"
f./gene="1gtA"

CDS
106..1152
f./gene="1gtA"

/allele="6"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtA"
/protein_id="AA033869.1"
/db_xref="GI:21069196"
/translation="MPSERFRHRRAVREKLOPLYSVLICATNVEKYPAGSLAAVNO
TWNLDILVDSSTDTGLAIADPOKRSRIKILAOONSGILPELNLGLBLASG
GSEYIARTADIAIPADWTEKIVGEMKDSRIIANGAMLEVLSEBKDNRLAHHHG
KIMKPTREDIADPEPFGNPIHNNMTIMRSEVLDGLVNTERRDWAEOYQFYDVSK
LRLAAYPEALVKYRLHANOVSKSVRCHIAOGIOKTRARDPLQSMGPKRPFSLB
YROIYAVALELLEKHLPEBDFBARARFLYQCRKRTDTLPAGAMLPDAADGRRLFTL
ROYGILRHLLNKR"
1194..2033
/gene="1gtB"
/allele="7"
1194..2033
/gene="1gtB"
/allele="7"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtB"
/protein_id="AA033870.1"
/db_xref="GI:21069197"
/translation="MKNHVISLASAERRAHIAITFGARGLPPOFPDMLMPSERLQA
MAELVGLSAPHYLSGVKACFMASHAVLMKQALDGLPIAVFEDVILGSEAKELA
EDAMLOERFDKDSAFIVLETFMFMHVLTPSGVDYGGRAFLPILSEHMGTAGYIIS
KAMRPFILRPALPPEGHLPVDMWSPFDEBGMVCOIDPALCAOBLHVKFPHON
SALGSLIHDRILNKKQRDSPATPFHRLIRALTKISREBKRRORRQULGKIIV
PQO"
2182..2988
/gene="1gtH"
/allele="7"
2182..2988
/gene="1gtH"
/allele="7"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtH"
/protein_id="AA033871.1"
/db_xref="GI:21069198"
/translation="MKNHVISLASAERRAHIAITFGARGLPPOFPDMLMPSERLQA
MAELVGLSAPHYLSGVKACFMASHAVLMKQALDGLPIAVFEDVILGSEAKELA
EDAMLOERFDKDSAFIVLETFMFMHVLTPSGVDYGGRAFLPILSEHMGTAGYIIS
KAMRPFILRPALPPEGHLPVDMWSPFDEBGMVCOIDPALCAOBLHVKFPHON
SALGSLIHDRILNKKQRDSPATPFHRLIRALTKISREBKRRORRQULGKIIV
PQO"
93.7% Score 775.6; DB 1; Length 3289;
Best Local Similarity 97.1%; Pred. No. 2.9e-174;
Matches 790; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGCAAAACACGTTATCACTTACCTCCGCGCAAGACGAGGCGCACATTGCCAT 60
Db 1194 ATGCAAAACACGTTATCACTTACCTCCGCGCAAGACGAGGCGCACATTGCCAT 1253
Qy 61 ACCTTCGAGGACAGGATCCCGTTTTCAGAGCACTGATGCGGCTGTAAGG 120
Db 1254 ACCTTCGAGGACAGGATCCCGTTTTCAGAGCACTGATGCGGCTGTAAGG 1313
Qy 121 CTGGAACAGGACATGCGCAACTGCTCCCGGCTGTGCGGACACCCCTATTGAGCGGA 180
Db 1314 CTGGAACAGGACATGCGCAACTGCTCCCGGCTGTGCGGACACCCCTATTGAGCGGA 1373
Qy 181 GTGAAAAAGCCTGTTTATGAGCCAGCGCGTATTTGGAAGCAGGCAATTGAGCAAGCT 240
Db 1374 GTGAAAAAGCCTGTTTATGAGCCAGCGCGTATTTGGAAGCAGGCAATTGAGCAAGCT 1433
Qy 241 CTGCGGTATACCGGTATTTGAGGACGAGTTTACTCGCGCAAGGAGGAAAAATTC 300

Db 1434 CTGCCGTATATGCGCGTATTTGAGAGCAGCGTTTTCCTGCGGAGAGCGCGGAGAAATTC 1493
Qy 301 CTTCGCGAAGACGCTGGCTGCAAGACGCTTTGACCGGATACCGCTTTATCGTCCG 360
Db 1494 CTTCGCGAAGACGCTGGCTGCAAGAGCGTTTTCGACAGAAATTCCTCTTTATCGTCCG 1553
Qy 361 TTGGAACAGATGTTTATGACGCTGACCTGCGCTCCGGGCTGGCGGATTAATCGCGGG 420
Db 1554 TTGGAACAGATGTTTATGACGCTGACCTGCGCTCCGGGCTGGCGGATTAATCGCGGG 1613
Qy 421 CGGCGCTTTCCGCTGTTGGAAGACGAACTGCGGGAGCGCGGCTATATCAATTTCCGA 480
Db 1614 CGGCGCTTTCCGCTGTTGGAAGACGAACTGCGGGAGCGCGGCTATATCAATTTCCGA 1673
Qy 481 AAGAGCATGCGGTTTTCCTGGAAGGTTTTCGCGCTCCGCGCGGAGAGGCTGACCGCC 540
Db 1674 AAGAGCATGCGGTTTTCCTGGAAGGTTTTCGCGCTCCGCGCGGAGAGGCTGACCGCC 1733
Qy 541 GTGCAATGATGATGTTTTCGAGCGATTTTTCGACAGGAGGAATGCGGTTTTCGAGCTC 600
Db 1734 GTGCAATGATGATGTTTTCGAGCGATTTTTCGACAGGAGGAATGCGGTTTTCGAGCTC 1793
Qy 601 AATCCGCGCTTGGGCGCGCAAGAGCTGCAATTAATGCAAGTTTTCGACGCAAAACGCGCA 660
Db 1794 GATCCGCGCTTGGGCGCGCAAGAGCTGCAATTAATGCAAGTTTTCGACGCAAAACGCGCA 1853
Qy 661 TTGGGAGAGCTGATTCGACAGCAGCGCTCTGGAACCGGAAACAGCAAGGCGGATTC 720
Db 1854 TTGGGAGAGCTGATTCGACAGCAGCGCTCTGGAACCGGAAACAGCAAGGCGGATTC 1913
Qy 721 CCGCGCAACATTCGAAACGCGCTGATCCGCGCTTGAACCAATATCGAGGAGAAAG 780
Db 1914 CCGCGCAACATTCGAAACGCGCTGATCCGCGCTTGAACCAATATCGAGGAGAAAG 1973
Qy 781 GAAAAACGCGCGCAAGGCGCGCAAGCTTCATTG 814
Db 1974 GAAAAACGCGCGCAAGGCGCGCAAGCTTCATTG 2007

RESULT 12
AP470655
LOCUS AP470655 3383 bp DNA linear BCT 22-MAY-2002
DEFINITION Neisseria meningitidis strain 35E 1gt-1 locus, complete sequence.
ACCESSION AP470655
VERSION AP470655.1 GI:21069168
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 3383)
Zhu, P., Klutch, M. J., Baah, M. C., Tsang, R. S., Ng, L. K. and Tsai, C. M.
Genetic diversity of three 1gt loci for biosynthesis of
lipooligosaccharide (LOS) in Neisseria species
Microbiology 148 (Pt 6), 1833-1844 (2002)
JOURNAL
PUBMED
12055304
2 (bases 1 to 3383)
Zhu, P.
Direct Submission
Submitted (22-JAN-2002) DBPAP, CBBR/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
location/Qualifiers
1..3383
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="35E"
/db_xref="taxon:487"
140..1186
/gene="1gtA"
/allele="16"
140..1186
/gene="1gtA"
/allele="16"
140..1186
/gene="1gtA"
/allele="16"

gene
CDS

/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtA"
/protein_id="AA033849.1"
/db_xref="gi:21069168"
/translation="MPSRAFRHRAVRENKQPIVSVLICAYNEKIPAGSLAVVNO
TWCNDILITVDGSTDGTLTAQRFQBOGRIKITLAQONSGLIPSLTGLDELAKSG
MGEYIARTDADIDIAAPDWITEKI VGEKEDRSI IAMGAWLEYSBKGNRLARHHEG
KIWKPPRHEIDIAEFPPGNDIHNNTIMERSVLDGRLRYNTERDMEDYQFWYDVK
LGRLLAYPEALVKTRLHAMOVSKYSIRIOMEIAQGI OKTANDEFLQSGWFTKPSLSL
YROIKAAYVELLEKHLPEDEPERARFLVQCFKRTDLPAGVWIDFANGMRRLFTL
ROYGILHRLKRN"
128..2067
/gene="1gtB"
/allele="4"
128..2067
/gene="1gtB"
/allele="4"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtB"
/protein_id="AA033850.1"
/db_xref="gi:21069170"
/translation="MONHVISLAAABERBAHIATPGRHGIPOFPDAMPSERLBOA
MAELVPGLSAPYLSGVKACFMGHVVLAKMDLDEGLPYLAVFEDDVLGGAQPLA
EDTWLEBEPDPSAFIVRLFTWPAKVI VREPKNLYENRSFPLESEHWGAGTISR
KAMRFLDERFALVPAEWIKAVDMMFTFYFDKEMPYVQNPALCTQELHAYAKFLSKN
SALGSLIEHRLRLRKQQRDS PANTKRLIRALTYKISRERKRRQREOLJCKITIV
PFO"
2216..3058
/gene="1gtE"
/allele="4"
2216..3058
/gene="1gtE"
/allele="4"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtE"
/protein_id="AA033851.1"
/db_xref="gi:21069171"
/translation="MONHITSLAABERBAHIATPGRHGIPOFPDAMPSERLBOA
MAELVPGLSAPYLSGVKACFMGHVVLAKMDLDEGLPYLAVFEDDVLGGAQPLA
EDTWLEBEPDPSAFIVRLFTWPAKVI VREPKNLYENRSFPLESEHWGAGTISR
KAMRFLDERFALVPAEWIKAVDMMFTFYFDKEMPYVQNPALCTQELHAYAKFLSKN
SALGSLIEHRLRLRKQQRDS PANTKRLIRALTYKISRERKRRQREOLJCKITIV
ISFR"
ORIGIN
Query Match 91.0%; Score 753.2; DB 1; Length 3383;
Best Local Similarity 95.3%; Pred. No. 6.4e-169;
Matches 776; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 1 ATGCAAAACGAGTATGAGCTTAGCTTCGCGCAGAGCAGAGGCGCATTGCCGAT 60
Db 1228 ATGCAAAACGAGTATGAGCTTAGCTTCGCGCAGAGCAGAGGCGCATTGCCGCA 1287
Qy 61 ACCTTCGAGAGCAGGATCCGTTTCACTTTTTCGACGCACTGATGCGCTCGAAAG 120
Db 1288 ACCTTCGAGAGCAGGATCCGTTTCACTTTTTCGACGCACTGATGCGCTCGAAAG 1347
Qy 121 CTGGAACAGGCAATGAGCGAATCGTCCCGCTTGTTCGCGCACCCCTATTGAGCGGA 180
Db 1348 CTGGAACAGGCAATGAGCGAATCGTACCGCGCTTGTTCGCGCACCCCTATTGAGCGGA 1407
Qy 181 GTGGAAGAAAGCTCTTTATGAGCGACCGCTATTTGGAAGCAGGCAATTGAGAGAGT 240
Db 1408 GTGGAAGAAAGCTCTTTATGAGCGACGTCATTTGGAAGCAGGCAATTGAGAGAGT 1467
Qy 241 CTGCCGTATATCAGCGTATTTGAGAGCAGCTTTTACTCGCGAAGGTGAGAGAAATTC 300

```

Db      1468 CTGCCGTATATCCCGTATTGAGAGCAGCTTTTACTCTCGCGAAGCGCGAGCAGTTC 1527
QY      301 CTTCGGAAGACGCTTGCGTGCAGAAACGCTTTGACCCGGGATACCGGCTTATCGTCCG 360
Db      1528 CTTCGGAAGACGCTTGCGTGCAGAAACGCTTTGACCCGGGATACCGGCTTATCGTCCG 1587
QY      361 TTGAAAACGATGTTTATGACAGTCTGACCTCGCCCTCGCGCGTGGCGATTATCTCGCG 420
Db      1588 TTGAAAACGATGTTTATGACAGTCTGACCTCGCCCTCGCGCGTGGCGAAATCTCGCG 1647
QY      421 CGGCGCTTTCCGCTGTTTGAAGAACGAAACACTGGGGGAGCGCGGCTATATCATTTCCGA 480
Db      1648 CGGCGCTTTCCGCTGTTTGAAGAACGAAACACTGGGGGAGCGCGCTATATCATTTCCAA 1707
QY      481 AAGCGATGCGGCTTTTCTGAGCAGGTTTGGCGCTCGCGCGGCGAAGGCGTGCACCC 540
Db      1708 AAGCGATGCGGCTTTTCTGAGCAGGTTTGGCGCTCGCGCGGCGAAGGCGTGCACCC 1767
QY      541 GTTCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db      1768 GTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1827
QY      601 AATCCCGCTTTGCGCGCCCAAGAGCTGCATTATGCAAGTTTCAACACCAAAACAGCGCA 660
Db      1828 AATCCCGCTTTGCGCGCCCAAGAGCTGCATTATGCAAGTTTCAACACCAAAACAGCGCA 1887
QY      661 TTGGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db      1888 TTGGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1947
QY      721 CCGCGCAACATTCATCAACACCGCTGCATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db      1948 CCGCGCAACATTCATCAACACCGCTGCATCCGCGCTTGAACCAAAATCAGCAGGAAAG 2007
QY      781 GAAAAACGCGCGCAAGCGCGCAACGCTTCATTG 814
Db      2008 GAAAAACGCGCGCAAGCGCGCAACGCTTCATTG 2041

RESULT 13
AF470665      3259 bp      DNA      linear      BCT 22-MAY-2002
DEFINITION   Neisseria meningitidis strain 7897 1gt-1 locus, complete sequence.
ACCESSION    AF470665
VERSION      AF470665.1 GI:21069204
KEYWORDS
SOURCE
ORGANISM     Neisseria meningitidis
              Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
              Neisseriaceae; Neisseria.
REFERENCE
AUTHORS      Zhu, P., Klutch, M.J., Bash, M.C., Tsang, R.S., Ng, L.K. and Tsai, C.M.
TITLE        Genetic diversity of three 1gt loci for biosynthesis of
              lipooligosaccharide (LOS) in Neisseria species
JOURNAL      Microbiology 148 (Pt 6), 1833-1844 (2002)
PUBMED       12055303
REFERENCE
AUTHORS      Zhu, P.
TITLE        Direct Submission
JOURNAL      Submitted (22-JAN-2002) DBPAP, CBER/FDA, 8800 Rockville Pike,
              Bethesda, MD 20892, USA
FEATURES
source       1..3259
              location/Qualifiers
              /organism="Neisseria meningitidis"
              /mol_type="Genomic DNA"
              /strain="7897"
              /db_xref="taxon:487"
              106..1194
              /gene="1gtA"
              /allele="11"
              106..1194
              /gene="1gtA"
              /allele="11"
              /allele="11"

CDS

```

```

/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtA"
/protein_id="AAM33875.1"
/db_xref="GI:21069205"
/translation="MPEAFRRHRAVRENKLOPLVSLICAVNVEKYPASOLAAYNO
TWRLIELIVDSDTDTLAIORFOBOGRIRILAOPRNSGLISLIGLELAKSG
GGEYIARTDADIDIAIPDMTEKIVGEMEDRSIIMAGMIEVISEEDGRIARHRHG
KIMKQTRHEDIDAFPPGNPIHNTMTWRSVITOGELNTERMDADYOFYDVSK
KLRLAYYPEALVKYRLHANQVSSKYSIRHETAOGIQKTAENDPLQSMFKTRFDSLE
YROIIVAYVELLEKILPEEDFERARFIYQCFKRTDTPAGAMLDFADGKRRRLFTM
ROYFILHRLIKRRQARSDSAGKEQT"
1194..2033
/gene="1gtB"
/allele="6"
1194..2033
/gene="1gtB"
/allele="6"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtB"
/protein_id="AAM33876.1"
/db_xref="GI:21069206"
/translation="MWNHVISLASHARRRAHTDPRGRTIPQGFDDLMPSERLEQA
MAELVPGISAPHYLSGVKACFMSHAIVLMKQLDGLPYIAFEDVILGSAEKFLA
EDAMIQERFPDPSAFIVRLTFMHLVTPSGVADYGRAPFLSBRHGMAGYIISR
KAMFFFLDRFAPVLPSEIRLKAVDWMLFSSFLDKGWTVCQLPALCVQSEITLPSQKNG
RSYRYNRRSPVILKRALGKIGREIRBARERKQKLEKILGRHVPE"
2182..2988
/gene="1gtH"
/allele="3"
2182..2988
/gene="1gtH"
/allele="3"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtH"
/protein_id="AAM33877.1"
/db_xref="GI:21069207"
/translation="MWNHVISLASHARRRAHTDPRGRTIPQGFDDLMPSERLEQA
MAELVPGISAPHYLSGVKACFMSHAIVLMKQLDGLPYIAFEDVILGSAEKFLA
EDAMIQERFPDPSAFIVRLTFMHLVTPSGVADYGRAPFLSBRHGMAGYIISR
KAMFFFLDRFAPVLPSEIRLKAVDWMLFSSFLDKGWTVCQLPALCVQSEITLPSQKNG
RSYRYNRRSPVILKRALGKIGREIRBARERKQKLEKILGRHVPE"

ORIGIN
Query Match      90.6%; Score 750; DB 1; Length 3259;
Best Local Similarity 95.1%; Pred. No. 3.7e-168;
Matches 774; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY      1 ATGCAAAACCAAGTATGAGCTTACCTCCCGCGAAGACGAGGCGCACTTGCCCAT 60
Db      1194 ATGCAAAACCAAGTATGAGCTTACCTCCCGCGAAGACGAGGCGCACTTACCAT 1253
QY      61 ACCTTCGCGAGCAGCGATCCCGTTTCAAGTTTTCAGAGCACTGATCCCGCTGAAGG 120
Db      1254 ACCTTCGCGAGCGCGATCCCGTTTCAAGTTTTCAGAGCACTGATCCCGCTGAAGG 1313
QY      121 CTGGAACGCGCAATGCGGAACCTGTCCTCCGCTTGTGCGGCGACCCCTATTGACCGA 180
Db      1314 CTGGAACGCGCAATGCGGAACCTGTCCTCCGCTTGTGCGGCGACCCCTATTGACCGA 1373
QY      181 GTGAAAAAGCGCTGTTATGAGCGACGCGTATTTGGAAGCGCACTTGACGAAGGT 240
Db      1374 GTGAAAAAGCGCTGTTATGAGCGACGCGTATTTGGAAGCGCACTTGACGAAGGT 1433
QY      241 CTGCGGTATATCACCGTATTTGAGAGCAGCTTTTACTCGCGCAAGTGAGGAAATTC 300
Db      1434 CTGCGGTATATCGCGGTATTTGAGAGTGAATGATCTCTTCGCGCGAAGCGCGGAGAAATTC 1493

```

QY 301 CTGCGAAGACGCTTGCTGCAAGAACGCTTTGACCCGGATACCGCTTTATGCTCCG 360
Db 1494 CTGGCGGAAGACGCTTGCTGCAAGAACGCTTTGACCCGGATACCGCTTTATGCTCCG 1553
QY 361 TTGGAAACGATGTTTATGCAAGTCTGCAAGCTCCGCTCCGCGTGGCGGATTACTCGGG 420
Db 1554 TTGGAAACGATGTTTATGCAAGTCTGCAAGCTCCGCTCCGCGTGGCGGATTACTCGGG 1613
QY 421 CGCGCTTTCCGCTGTTGGAAAGCAACCTGGGGGACGGCGGGCTATATCATTTCCCA 480
Db 1614 CGGCTTTCCGCTGTTGGAAAGCAACCTGGGGGACGGCGGGCTATATCATTTCCCA 1673
QY 481 AAAGCATGCGGCTTTTCTGCAAGGTTTCCGCTCCGCGCGCAAGGCTGCAAGCC 540
Db 1674 AAAGCATGCGGCTTTTCTGCAAGGTTTCCGCTCCGCGCGCTGCAAGGCTGCAAGCC 1733
QY 541 GTGCAATGATGATGTTTCAGCCGATTTTTCGACAGGAAAGATGCGGTTTCCGAGCTC 600
Db 1734 GTGCAATGATGATGTTTCAGCCGATTTTTCGACAGGAAAGATGCGGTTTCCGAGCTC 1793
QY 601 AATCCCGCTTGCTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660
Db 1794 AATCCCGCTTGCTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 1853
QY 661 TTGGGAGAGCTGATGCAACAGACCGCTCTGCAACCGCAAAAGCGCGATTC 720
Db 1854 TTGGGAGAGCTGATGCAACAGACCGCTCTGCAACCGCAAAAGCGCGATTC 1913
QY 721 CCGCGCAACATTCGAAACAGCGCTGATCCGCGCTTGCACCAATTCAGCAGGAAAG 780
Db 1914 CCGCGCAACATTCGAAACAGCGCTGATCCGCGCTTGCACCAATTCAGCAGGAAAG 1973
QY 781 GAAAAACGCGCGCAAGGCGCGCAACGTTCAATG 814
Db 1974 GAAAAACGCTCGCGCAAGGCGCGCAACGTTCAATG 2007

RESULT 14
LOCUS AY134878 3066 bp DNA linear BCT 02-AUG-2003
DEFINITION Neisseria polysaccharea strain 87043 lgt-1 locus, complete
ACCESSION AY134878
VERSION AY134878.1 GI:33413664
KEYWORDS
SOURCE Neisseria polysaccharea
ORGANISM Neisseria polysaccharea
REFERENCE 1 (bases 1 to 3066)
AUTHORS Zhu, P. and Tsai, C.M.
TITLE Galactosyltransferase gene family in Neisseria: Implication for understanding the evolution of prokaryotic enzymes responsible for glycosylation
JOURNAL Unpublished
REFERENCES 2 (bases 1 to 3066)
AUTHORS Zhu, P.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2002) LBP/DBPAP/CBER, FDA, 8800 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source location/Qualifiers
1. 3066
/organism="Neisseria polysaccharea"
/mol_type="genomic DNA"
/db_xref="taxon:489"
/note="strain 87043"
181..1230
/gene="lgtA"
181..1230
/allele="19"
181..1230
/gene="lgtA"
/allele="19"

CDS
181..1230
/gene="lgtA"
/allele="19"

/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="lgtA"
/protein_id="AA08517.1"
/db_xref="GI:33413665"
/translation="MPSEAPRRHRAVRENKLOPIVSVLICAVNEKYPAGSLAAVNO
TWRNDILIVDGGTDGTLTAORPOEDDGRILIAORNSGLIPSLNIGDELAKSG
GGGELIARTDADDDIAAPDWIKETIGEMKESIIAMGMLVLSERKGNLARHEH
GKIWKQPRHEDIAAPFPNGPPIHNNMTKRSYVDGSLRNTEREDMAEDYQPTDVS
KUGRLATYPEALVKTIRLANOVSSKTSYRQHEIAQGIQKTANPFLQSMGKRTFDSL
EYRQIKAVAYBELRKHLPEDFERARRELQCFKRTDTPAGAMLDFADMRMLFT
LROYGLIRLRLLKRR"
1272..2111
/gene="lgtB"
/allele="19"
1272..2111
/gene="lgtB"
/allele="19"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="lgtB"
/protein_id="AA08518.1"
/db_xref="GI:33413666"
/translation="MONHVISLAAABERRAHIAITFGRHGIPQFPDAMPSELEQA
MAELVPGISAHPIYLSGVKACFMGHTVLMKQALDESPLYIAVFEDVVLGGAEKFLA
EDGMKERPDSDAFIVLETFMNVHLTSPGVADYGRAPPLLESBHFGAGYIISR
KAMRFLDRFALPSEGLPYDMMVFGDDRBGMPCVOLPALCAOELHAYKHPDON
SALGSLIHDLRLNRKQORSDSPANTFPHRLIRLTLKISREKRRQRELDIGKITV
PQO"
2260..3066
/gene="lgtH"
/allele="12"
2260..3066
/gene="lgtH"
/allele="12"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="lgtH"
/protein_id="AA08519.1"
/db_xref="GI:33413667"
/translation="MONHVISLAAABERRAHIAITFGRHGIPQFPDAMPSELEQA
MAELVPGISAHPIYLSGVKACFMGHTVLMKQALDESPLYIAVFEDVVLGGAEKFLA
EDTMLEKFPDSDAFIVLETFMNVHLTSPGVADYGRAPPLLESBHFGAGYIISR
KAMRFLDRFALPSEGLPYDMMVFGDDRBGMPCVOLPALCAOELHAYKHPDON
RQESYRNRSFKVLKRALGKIGREIERARERKQKLEKLGHVYFE"

ORIGIN
Query Match 90.4%; Score 748.4; DB 1; Length 3066;
Best Local Similarity 95.0%; Pred. No. 9e-168;
Matches 773; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 ATGCAAAACGAGTTATGAGCTTCCGCGCAGAACGAGGCGCACTATGCGCAT 60
Db 1272 ATGCAAAACGAGTTATGAGCTTCCGCGCAGAACGAGGCGCACTATGCGCAT 1331
QY 61 ACCTTCGAGGAGCAGGATCCCGTTGAGTTTTCGACGCACTGATGCGCTGAAAG 120
Db 1332 ACCTTCGAGGAGCAGGATCCCGTTGAGTTTTCGACGCACTGATGCGCTGAAAG 1391
QY 121 CTGGAACAGGCAATGCGGAGACTCGTCCCGGCTTGTGGGCGACCCCTATTTGAGCGGA 180
Db 1392 CTGGAACAGGCAATGCGGAGACTCGTCCCGGCTTGTGGGCGACCCCTATTTGAGCGGA 1451
QY 181 GTGAAAAAGCTGCTTTATGAGCCAGCGCTATTTGGAAGCAGGCAATTGACGAAGGT 240
Db 1452 GTGAAAAAGCTGCTTTATGAGCCAGCACCGTATTTGGAAGCAGGCGCTGACGAAGGC 1511
QY 241 CTGCGGTATACCGATTTGAGAGCAGCTTTTACCTCGGCGAAGGAGGAAAAATTC 300
Db 1512 TTACCGTATTTGCGGATTTGAGAGCAGCTTTTACCTCGGCGAAGGCGGAAAAATTC 1571

QY	301	CTTGGCGAAGAGCGCTTGCGTGCAAAAGCGCTTGAACCGGAAATACCGCTTTATGCTCCG	360
Db	1572	CTTGGCGAAGAGCGCTTGCGTGCAAAAGCGGTTGACCGGATTCGGCTTTATGCTTCGT	1631
QY	361	TTGGAACGATGTTTATGACGCTCTGACCTTCGCGCTTCGCGGATTAATCTGCGAG	420
Db	1632	TTGGAACGATGTTTATGACGCTCTGACCTTCGCGCTTCGCGGATTAATCTGCGAG	1691
QY	421	CGCGGCTTTCGGCTGTGTGAAAGCGAACAATGCGGGAAGCGCGGCTTATATCAATTTCCCGA	480
Db	1692	CGTGGCTTTCGGCTGTGTGAAAGCGAACAATGCGGGAAGCGCGGCTTATATCAATTTCCCGA	1751
QY	481	AAAGGATGCGGTTTCTTCGTGACAGTTTTCGCGCTTCGCGCGCGCGAAGGCGTGACCCC	540
Db	1752	AAAGGATGCGGTTTCTTCGTGACAGTTTTCGCGCTTCGCGCGCGCGAAGGCGTGACCCC	1811
QY	541	GTGCACTGATGATGTTTTCGACGATTTTTCGACAGGAAAGAAATGCGGTTTTCGACGCTC	600
Db	1812	GTGCACTGATGATGTTTTCGACGATTTTTCGACGAGAAAGAAATGCGGTTTTCGACGCTC	1871
QY	601	AATCCCGCTTGTGGCGCCCAAGGCTGCTTATGCGCAATTTACAGACCAAAACAGCGCA	660
Db	1872	AATCCCGCTTGTGGCGCCCAAGGCTGCTTATGCGCAATTTACAGACCAAAACAGCGCA	1931
QY	661	TTGGGACGCTGATGCAACAGACCGCGCTTCGGAACGCGAACAAGCGGAGCGCATTC	720
Db	1932	TTGGGACGCTGATGCAACAGACCGCGCTTCGGAACGCGAACAAGCGGAGCGCATTC	1991
QY	721	CCCGCGAACACATTGAAACACCGCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG	780
Db	1992	CCCGCGAACACATTGAAACACCGCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG	2051
QY	781	GAAAGCGCGCGCAAGGCGCGAACAAGTTCAATG	814
Db	2052	GAAAGCGCGCGCAAGGCGCGAACAAGTTCAATG	2085
RESULT 15			
LOCUS	NMA22491	349061 bp	DNA linear BCT 16-APR-2005
DEFINITION	Neisseria meningitidis serogroup A strain Z2491 complete genome; segment 2/7.		
ACCESSION	AL162753	AL157959	
VERSION	AL162753.2	GI:7379120	
KEYWORDS			
SOURCE	Neisseria meningitidis Z2491		
ORGANISM	Neisseria meningitidis Z2491		
REFERENCE	Neisseria meningitidis Z2491		
AUTHORS	Neisseria meningitidis Z2491		
	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.		
	1 (bases 1 to 349061)		
	Parkhill, J., Achtman, M., James, K. D., Bentley, S. D., Churcher, C., Kite, S. R., Morelli, G., Basham, D., Brown, D., Chillingworth, T., Davies, R. M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N., Holtroyd, S., Jagsels, K., Leather, S., Moule, S., Mungall, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Simmonds, M., Skellern, J., Whitehead, S., Spratt, B. G. and Barrell, B. G.		
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491		
JOURNAL	Nature	404 (6777),	502-506 (2000)
PUBMED	10761919		
REFERENCE	2 (bases 1 to 349061)		
AUTHORS	Parkhill, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-2000)		
COMMENT	Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
NOTES	Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web.		
FEATURES	(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).		
SOURCE	Location/Qualifiers		
	1. 349061		

```

repeat_unit
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/notes="serogroup: A"
complement(24..206)
/notes="ATR repeat; hmms hit to HMM ATR (1 - 183), score:
310.39"
/label=ATR
RBS
209..212
220..681
CDS
/gene="NMA0368"
220..681
/notes="NMA0368"
153 aa; contains four probable transmembrane domains"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB83669.1"
/db_xref="GI:7379121"
/db_xref="UniProt/TREMBL:Q9JWH7"
/translation="MOEONRKSFPFIVMLVLAIVLAIKRVVPAVAFWGLALVAFSKAVYV
LIISFSLIDIRYCAVYANYVMAIVLAIKRVVPAVAFWGLALVAFSKAVYV
DEAGTISIVRYGAGFIIWAAFAVVASIGTFPAGNKKRKAASADGTNDY"
671..1492
/gene="hemK"
671..1492
CDS
/gene="hemK"
/notes="NMA0369, hemK, Hemk protein, len: 273 aa; similar
to e.g. HEK, ECOLI P37186 HEK protein (277 aa), fasta
scores; E(): 0, 42.3% identity in 279 aa overlap, Contas
PS00092 N-6 Adenine-specific DNA methylase signature"
/codon_start=1
/transl_table=11
/product="hemK protein"
/protein_id="CAB83670.1"
/db_xref="GI:7379122"
/db_xref="GOA:Q9JWH6"
/db_xref="InterPro:IPR002052"
/db_xref="InterPro:IPR004556"
/db_xref="InterPro:IPR004556"
/db_xref="UniProt/TREMBL:Q9JWH6"
/translacion="MTFDKMLGSLKLPKKNRMLLOVSEYTRVOLLTRGSEEMPEV
RQARDKRLINSEPAVYILGAREFGRTVNPSTVILPRPTHEHVEAVILRLPEN
GKRVMLDGLGSAVAIVTALEPDPFAFVASISIPPALETAKNADIGARVEFAYGSMF
DTDSREKMKIIVSNPVIENGKHLQSGDLAREPDIATLDSDDLSCRTLAQAP
DLRAEGGTLILHEGPDGAAVRGVLAENGFSVETLLPDLGLDRLVTLGKTMKLK"
831..840
/misc_feature
/gene="hemK"
/notes="Core DNA uptake sequence: gccgctcgaa"
/label=DUS
1162..1171
/misc_feature
/gene="hemK"
/label=DUS
/notes="Core DNA uptake sequence: gccgctcgaa"
1187..1207
/misc_feature
/gene="hemK"
/notes="PS00092 N-6 Adenine-specific DNA methylases
signature"
complement(1279..1288)
/misc_feature
/label=DUS
/notes="Core DNA uptake sequence: gccgctcgaa"
1565..1568
1578..12966
CDS
/gene="NMA0370"
1578..12966
/gene="NMA0370"
/notes="NMA0370", probable integral membrane protein, len:
462 aa; similar to hypothetical proteins e.g. Y325_HARIN
P44630 hypothetical protein H10325 (450 aa), fasta_scores;
E(): 0, 49.8% identity in 464 aa overlap"
/codon_start=1

```


QY 721 CCGGCAACATTCACAACCCGCTGATCGGCTTGA CCAAAATCAGCAGGAAAG 780
Db 163504 CCGGCAACATTCACAACCCGCTGATCGGCTTGA CCAAAATCAGCAGGAAAG 163563
QY 781 GAAAAAGCCGCGCAAGGCGCAACATTCATTG 814
Db 163564 GAAAAAGTCGCGCAAGGCGCAACATTCATTG 163597

Search completed: April 7, 2006, 18:04:58
Job time : 5663.11 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:06 ; Search time 933.991 Seconds
(without alignments)
5908.370 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacacagctatcag.....tcattgtgccttccataa 828

Scoring table: IDENTITY_NUC
Gapox 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333246308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	828	2	AAx84281 N. mening
2	826.4	99.8	46593	3	AAx81456 N. mening
3	826.4	99.8	110000	3	AAx81456 N. mening
4	826.4	99.8	349980	3	AAx81456 N. mening
5	726	87.7	5859	2	AAx81456 N. mening
6	726	87.7	5859	2	AAx81456 N. mening
7	721.2	87.1	837	10	ABx38360 N. gonorr
8	471.8	57.0	840	10	ABx38360 N. gonorr
9	426	51.4	840	8	ABx38360 N. gonorr
10	93	11.2	14547	13	ADx09919 N. mening
11	93	11.2	10645	13	ADx09919 N. mening
12	86.6	10.5	110000	2	AAx81456 N. mening
13	73.2	8.8	768	12	ADx09919 N. mening
14	73.2	8.8	4583	14	ADx09919 N. mening
15	73.2	8.8	5584	14	ADx09919 N. mening
16	73.2	8.8	65792	4	AAx81456 N. mening
17	71.4	6.9	5570	14	ADx09919 N. mening
18	57	6.9	3625	13	ADx09919 N. mening
19	57	6.9	110000	2	AAx81456 N. mening

20	57	6.9	135356	13	ADx05646	Adx05646 Haemophil
21	48.2	5.8	2000	8	ADx05646	Adx05646 Haemophil
22	47	5.7	2000	8	ADx05646	Adx05646 Haemophil
23	39.8	4.8	4532	4	ABx10847	ABx10847 Drosophila
24	39.4	4.8	4455	14	ACx170656	ACx170656 M. xanthu
25	39.4	4.8	4455	14	ACx170656	ACx170656 M. xanthu
26	39	4.7	399	14	ACx166474	ACx166474 M. xanthu
27	39	4.7	984	4	ABx125049	ABx125049 Drosophila
28	39	4.7	3068	4	ABx125049	ABx125049 Drosophila
29	39	4.7	23417	14	ACx164744	ACx164744 M. xanthu
30	38.6	4.7	753	11	ABx11686	ABx11686 Pseudomon
31	38.6	4.7	954	13	ADx42128	ADx42128 Bacteri
32	38.6	4.7	1260	11	ABx11889	ABx11889 Pseudomon
33	38.6	4.7	1635	11	ABx11889	ABx11889 Pseudomon
34	38.6	4.7	1743	11	ABx12070	ABx12070 Pseudomon
35	38	4.6	1760	14	ABx26876	ABx26876 P. rad
36	37.8	4.6	2611	6	ABx56602	ABx56602 S. macro
37	37.8	4.6	4540	6	ABx56602	ABx56602 S. macro
38	37.4	4.5	1329	13	ADx20087	ADx20087 Plant
39	36.8	4.4	31422	3	AAx92302	AAx92302 S. averm
40	36.8	4.4	31422	4	AAx92302	AAx92302 S. averm
41	36.8	4.4	39912	14	ADx43464	ADx43464 S. averm
42	36.4	4.4	1122	8	ACA44178	ACA44178 Prokaryot
43	36.4	4.4	7788	10	ADx23894	ADx23894 Streptomy
44	36.4	4.4	19211	3	AAx81507	AAx81507 N. mening
45	36.4	4.4	37360	10	ADx23892	ADx23892 Streptomy

ALIGNMENTS

RESULT 1	AAx84281	standard; DNA; 828 BP.
ID	AAx84281	
XX	AAx84281	
AC	AAx84281	
XX	AAx84281	
DT	08-SEP-1999	(first entry)
XX	08-SEP-1999	
DE	N. meningitidis Beta-1,4-galactosyltransferase coding sequence.	
XX	N. meningitidis Beta-1,4-galactosyltransferase coding sequence.	
KW	Beta-1,4-galactosyltransferase; accessory enzyme; catalytic domain;	
KW	glycosyltransferase; accessory enzyme; nucleotide sugar formation;	
KW	saccharide donor; oligosaccharide synthesis;	
KW	carbohydrate structure development; ds.	
OS	Neisseria meningitidis.	
XX	Neisseria meningitidis.	
PN	W09931224-A2.	
XX	W09931224-A2.	
PD	24-JUN-1999.	
XX	24-JUN-1999.	
PF	15-DEC-1998;	98WC-CA001180.
XX	15-DEC-1998;	
PR	15-DEC-1997;	97US-0069443P.
XX	15-DEC-1997;	
PR	14-DEC-1998;	98US-00211691.
XX	14-DEC-1998;	
PA	(CANADA) NAT RES COUNCIL CANADA.	
XX	(CANADA) NAT RES COUNCIL CANADA.	
PI	Gilbert M, Young NM, Wakarchuk WM;	
XX	Gilbert M, Young NM, Wakarchuk WM;	
DR	WPI, 1999-395174/33.	
XX	WPI, 1999-395174/33.	
PT	F-PSDB; AAY22156.	
XX	F-PSDB; AAY22156.	
PS	A new glycosyltransferase fusion protein useful in the enzymatic	
XX	synthesis of oligosaccharides.	
PS	Example 2; Fig 2; 63pp; English.	
XX	Example 2; Fig 2; 63pp; English.	
CC	This sequence encodes the Neisseria meningitidis Beta-1,4-	
CC	galactosyltransferase (also referred to as lgfb). The invention relates	
CC	to a nucleic acid encoding a fusion protein that comprises a	
CC	glycosyltransferase catalytic domain and a catalytic domain from an	

accessory enzyme that is involved in formation of a nucleotide sugar
which is a saccharide donor for a glycosyltransferase reaction. The
fusion protein is useful in the enzymatic synthesis of oligosaccharides.
The fusion proteins are able to catalyze more than one reaction involved
in the enzymatic synthesis. This is useful for the development of
therapeutic agents that have specific carbohydrate structures.
Carbohydrates are involved in recognition elements on the surface of
cells. The fusion protein can be used for the synthesis of both natural
carbohydrates and synthetic derivatives with novel properties. The fusion
polypeptide allows two glycosyltransferase reactions in a single vessel,
provides improved yields of end products. Additionally, cleanup and
disposal of extra solvents and by-products is reduced. The fusion protein
can also use directly different donor analogues and various acceptors
with a terminal galactose residue

Sequence 828 B; 190 A; 241 C; 224 G; 173 T; 0 U; 0 Other;

Query Match 100.0%; Score 828; DB 2; Length 828;

Best Local Similarity 100.0%; Pred. No. 4,1e-242; Mismatches 0; Gaps 0;

Matches 828; Conservative 0; Indels 0; Gaps 0;

```
QY 1 ATGCAAAACCGAGTATACGCTTAGCTCCGCGAGAAACGAGGCGACATTGCCGAT 60
DB 1 ATGCAAAACCGAGTATACGCTTAGCTCCGCGAGAAACGAGGCGACATTGCCGAT 60
QY 61 ACCCTCGGAGGACGCGCATCCCGTTTCAGTTTTCAGGCACTGATCCGCTGAAAG 120
DB 61 ACCCTCGGAGGACGCGCATCCCGTTTCAGTTTTCAGGCACTGATCCGCTGAAAG 120
QY 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTCGCGCAGACCCCTATTTCAGCGGA 180
DB 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTCGCGCAGACCCCTATTTCAGCGGA 180
QY 181 GTGGAAGAAAGCTCTTTATGAGCCAGCGCGTATTTGGAAGAGGCAATTGACGAAGT 240
DB 181 GTGGAAGAAAGCTCTTTATGAGCCAGCGCGTATTTGGAAGAGGCAATTGACGAAGT 240
QY 241 CTGCGGTATATACCGGTATTTGAGAGCAGCGTTTATCTGGGAGAGGTAGAGAAATTC 300
DB 241 CTGCGGTATATACCGGTATTTGAGAGCAGCGTTTATCTGGGAGAGGTAGAGAAATTC 300
QY 301 CTGCGGTATATACCGGTATTTGAGAGCAGCGTTTATCTGGGAGAGGTAGAGAAATTC 360
DB 301 CTGCGGTATATACCGGTATTTGAGAGCAGCGTTTATCTGGGAGAGGTAGAGAAATTC 360
QY 361 TTGGAAGAGATGTTTATGACGCTGACCTGCGCGTCCGCGGATTAATTCGCGG 420
DB 361 TTGGAAGAGATGTTTATGACGCTGACCTGCGCGTCCGCGGATTAATTCGCGG 420
QY 421 CGCGCTTTCCGCTGTTGGAAGAGCACTGCGGAGCGCGGCTATATCTTCCGA 480
DB 421 CGCGCTTTCCGCTGTTGGAAGAGCACTGCGGAGCGCGGCTATATCTTCCGA 480
QY 481 AAAAGGATCGGTTTCTCTGACAGGTTTCCGCGCTCCGCGGATTAATTCGCGG 540
DB 481 AAAAGGATCGGTTTCTCTGACAGGTTTCCGCGCTCCGCGGATTAATTCGCGG 540
QY 541 GTGCAATCTGATGATGTTTTCAGATTTTTCGACAGGAGAAATGCGGTTTCCAGCTC 600
DB 541 GTGCAATCTGATGATGTTTTCAGATTTTTCGACAGGAGAAATGCGGTTTCCAGCTC 600
QY 601 AATCCCGCTTGTGCGCCCAAGAGCTGATTAATTCGCAAGTTCACGACCAAAACAGCGCA 660
DB 601 AATCCCGCTTGTGCGCCCAAGAGCTGATTAATTCGCAAGTTCACGACCAAAACAGCGCA 660
QY 661 TTGGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 TTGGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 CCAGCAACACATTAACACCGCGCTGATCCGCGCTTTCAGCAAAATTCAGCAGGAGAAAG 780
DB 721 CCAGCAACACATTAACACCGCGCTGATCCGCGCTTTCAGCAAAATTCAGCAGGAGAAAG 780
```

QY 781 GAAAGACCGCGCAAGGCGGACGATTCATTGCTTCATTA 828
DB 781 GAAAGACCGCGCAAGGCGGACGATTCATTGCTTCATTA 828

RESULT 2

AAA81456/c
ID AAA81456 strand; DNA; 46593 BP.

AAA81456;

04-DEC-2000 (first entry)

N. meningitidis partial DNA sequence gnm_4 SEQ ID NO:4.

Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

antigen; vaccine; diagnosis; infection; antibacterial; identification;

Neisseria meningitidis.

MO200022430-AZ.

20-APR-2000.

08-OCT-1999; 99WO-US023573.

09-OCT-1998; 98US-0103794P.

30-APR-1999; 99US-0132068P.

(CHIR) CHIRON CORP.

Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V, Galeotti C, Mora M, Ratti G, Scarcelli M, Scarlato V, Rapuoli R, Pizza M;

WPI; 2000-318079/27.

Isolated nucleotide sequences of Neisseria meningitidis which can be used

in the diagnosis and treatment of N. meningitidis infection and other

Neisseria infections, for example, N.gonorrhoea.

Claim 7; Page 260-274; 1760dp; English.

The present invention describes methods of obtaining immunogenic proteins
from Neisseria genomic sequences. AAA81453 to AAA8244 represent
specifically claimed Neisseria meningitidis genomic DNA sequences;
AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
sequences and their corresponding proteins; AAA81254 to AAA81259 and
AAA81304 to AAA81321 represent PCR primers used in the isolation of
Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
used in the exemplification of the present invention. The nucleic acid
sequences, protein sequences, and antibodies against them, can be used in
the manufacture of a composition. The composition can be used as a
medicament (or in the manufacture of a medicament) for treating,
preventing or diagnosing infection due to Neisseria bacteria. For
example, some of the identified proteins could be components of vaccines
against Meningococcus B; against all serotypes; and/or against all
pathogenic Neisseriae. Identification of sequences from the bacterium
will also facilitate production of biological probes, particularly
organism-specific probes. Attempts to make efficacious Meningococcus B
vaccines have also been tried but none have successfully overcome
antigenic variability. The provision of further, complete sequences may
provide an opportunity to identify secreted or surface exposed proteins
that may be presumed targets for the immune system and which are not
antigenically variable or at least more conserved than other more
variable regions

Sequence 46593 B; 11355 A; 13195 C; 11355 G; 10687 T; 0 U; 1 Other;

Query Match 99.8%; Score 826.4; DB 3; Length 46593;

Mon Apr 10 08:45:27 2006

us-09-211-691-1.rng

Page 3

Best Local Similarity 99.9%; Pred. No. 7.1e-241;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCAAAACACGTTATACGTTAGCTTCCGCGCAGAAACGAGGCGCAATTGCCGAT 60
Db 33536 ATGCAAAACACGTTATACGTTAGCTTCCGCGCAGAAACGAGGCGCAATTGCCGAT 33477
QY 61 ACCTTGGCAGGCAACGCGATCCCGTTTCACTTTTTCAGACGACATGATCCGCTGAAAG 120
Db 33476 ACCTTGGCAGGCAACGCGATCCCGTTTCACTTTTTCAGACGACATGATCCGCTGAAAG 33417
QY 121 CTGGAACAGGCAATGCGGAATCGTCCCGGCTTGTGCGGCAACCCCTATTGAGCGGA 180
Db 33416 CTGGAACAGGCAATGCGGAATCGTCCCGGCTTGTGCGGCAACCCCTATTGAGCGGA 33357
QY 181 GTGGAAGAAAGCCCTGTTATGAGCAACGCGTATGTGGAAGCAAGGATTTGACGAAGT 240
Db 33356 GTGGAAGAAAGCCCTGTTATGAGCAACGCGTATGTGGAAGCAAGGATTTGACGAAGT 33297
QY 241 CTGCGGTATATCAACCGTATTGAGCAACGCTTTTACTCGCGAAGGTGAGAAATTC 300
Db 33296 CTGCGGTATATCAACCGTATTGAGCAACGCTTTTACTCGCGAAGGTGAGAAATTC 33237
QY 301 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCGCG 360
Db 33236 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCGCG 33177
QY 361 TTGGAAGACGATGTTATGACAGCTCTGACCTCGCGCTCGGCGGATTAATCGCGG 420
Db 33176 TTGGAAGACGATGTTATGACAGCTCTGACCTCGCGCTCGGCGGATTAATCGCGG 33117
QY 421 CGGCGCTTTCGCGTGTGGAAGACGCACTGGGGGACGCGGGGCTATATCATTTCCGA 480
Db 33116 CGGCGCTTTCGCGTGTGGAAGACGCACTGGGGGACGCGGGGCTATATCATTTCCGA 33057
QY 481 AAAGCATGCGGTTTCTCTGCAAGGTTTCCGCGCTGCGCGCAAGGCGTGCAACCC 540
Db 33056 AAAGCATGCGGTTTCTCTGCAAGGTTTCCGCGCTGCGCGCAAGGCGTGCAACCC 32997
QY 541 GTGCACTGATGATGTTGACGCAATTTTTCAGACAGGAGAAATGCGGTTTCCGACCT 600
Db 32996 GTGCACTGATGATGTTGACGCAATTTTTCAGACAGGAGAAATGCGGTTTCCGACCT 32937
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 32936 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 32877
QY 661 TTGGGAGCCTGATTCGACACGACCGCTCTGAAACGCAAAAGGCGCGATTC 720
Db 32876 TTGGGAGCCTGATTCGACACGACCGCTCTGAAACGCAAAAGGCGCGATTC 32817
QY 721 CCCGCAACACATTCAAAACGCGCTGATCGCGCTTTCAGCAAAATCAGACAGGAAAG 780
Db 32816 CCCGCAACACATTCAAAACGCGCTGATCGCGCTTTCAGCAAAATCAGACAGGAAAG 32757
QY 781 GAAAAACGCGGCAAGGCGCGCAAGCTTCATGTGCTCTTCCAAATTA 828
Db 32756 GAAAAACGCGGCAAGGCGCGCAAGCTTCATGTGCTCTTCCAAATTA 32709
```

RESULT 3

Continuation (6 of 9) of AAA81489 from base 500001 (N. meningitidis partial DNA sequence
WP Sequence split into 9 fragments LOCUS AAA81489 Accession AAA81489

Fragment Name	Begin	End
AAA81489_0	1	110000
AAA81489_1	100001	210000
AAA81489_2	200001	310000
AAA81489_3	300001	410000
AAA81489_4	400001	510000
AAA81489_5	500001	610000
AAA81489_6	600001	710000
AAA81489_7	700001	810000

WP AAA81489_8 800001 837096
Query Match 99.8%; Score 826.4; DB 3; Length 110000;
Best Local Similarity 99.9%; Pred. No. 1e-240;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCAAAACACGTTATACGTTAGCTTCCGCGCAGAAACGAGGCGCAATTGCCGAT 60
Db 91427 ATGCAAAACACGTTATACGTTAGCTTCCGCGCAGAAACGAGGCGCAATTGCCGAT 91368
QY 61 ACCTTGGCAGGCAACGCGATCCCGTTTCACTTTTTCAGACGACATGATCCGCTGAAAG 120
Db 91367 ACCTTGGCAGGCAACGCGATCCCGTTTCACTTTTTCAGACGACATGATCCGCTGAAAG 91308
QY 121 CTGGAACAGGCAATGCGGAATCGTCCCGGCTTGTGCGGCAACCCCTATTGAGCGGA 180
Db 91307 CTGGAACAGGCAATGCGGAATCGTCCCGGCTTGTGCGGCAACCCCTATTGAGCGGA 91248
QY 181 GTGGAAGAAAGCCCTGTTATGAGCAACGCGTATGTGGAAGCAAGGATTTGACGAAGT 240
Db 91247 GTGGAAGAAAGCCCTGTTATGAGCAACGCGTATGTGGAAGCAAGGATTTGACGAAGT 91188
QY 241 CTGCGGTATATCAACCGTATTGAGCAACGCTTTTACTCGCGAAGGTGAGAAATTC 300
Db 91187 CTGCGGTATATCAACCGTATTGAGCAACGCTTTTACTCGCGAAGGTGAGAAATTC 91128
QY 301 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCGCG 360
Db 91127 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCGCG 91068
QY 361 TTGGAAGACGATGTTATGACAGCTCTGACCTCGCGCTCGGCGGATTAATCGCGG 420
Db 91067 TTGGAAGACGATGTTATGACAGCTCTGACCTCGCGCTCGGCGGATTAATCGCGG 91008
QY 421 CGGCGCTTTCGCGTGTGGAAGACGCACTGGGGGACGCGGGGCTATATCATTTCCGA 480
Db 91007 CGGCGCTTTCGCGTGTGGAAGACGCACTGGGGGACGCGGGGCTATATCATTTCCGA 90948
QY 481 AAAGCATGCGGTTTCTCTGCAAGGTTTCCGCGCTGCGCGCAAGGCGTGCAACCC 540
Db 90947 AAAGCATGCGGTTTCTCTGCAAGGTTTCCGCGCTGCGCGCAAGGCGTGCAACCC 90888
QY 541 GTGCACTGATGATGTTGACGCAATTTTTCAGACAGGAGAAATGCGGTTTCCGACCT 600
Db 90887 GTGCACTGATGATGTTGACGCAATTTTTCAGACAGGAGAAATGCGGTTTCCGACCT 90828
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 90827 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 90768
QY 661 TTGGGAGCCTGATTCGACACGACCGCTCTGAAACGCAAAAGGCGCGATTC 720
Db 90767 TTGGGAGCCTGATTCGACACGACCGCTCTGAAACGCAAAAGGCGCGATTC 90708
QY 721 CCCGCAACACATTCAAAACGCGCTGATCGCGCTTTCAGCAAAATCAGACAGGAAAG 780
Db 90707 CCCGCAACACATTCAAAACGCGCTGATCGCGCTTTCAGCAAAATCAGACAGGAAAG 90648
QY 781 GAAAAACGCGGCAAGGCGCGCAAGCTTCATGTGCTCTTCCAAATTA 828
Db 90647 GAAAAACGCGGCAAGGCGCGCAAGCTTCATGTGCTCTTCCAAATTA 90600
```

RESULT 4

AAF21612/c
ID AAF21612 standard; DNA; 349980 BP.

ID	AAF21612	standard; DNA; 349980 BP.
AC	AAF21612;	
XX		
DT	13-MAR-2001 (first entry)	
XX		
DE	Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.	
XX		

KM Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;
OS Neisseria meningitidis.
XX MO200066791-A1.
XX 09-NOV-2000.
XX 08-MAR-2000; 2000MO-US005928.
XX 30-APR-1999; 99US-0132068P.
XX 08-OCT-1999; 99MO-US023573.
XX 28-FEB-2000; 2000GB-00004695.
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
PI Pizza M, Hickey E, Peterson J, Tetrelin H, Venter JC,
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,
PI Rapuoli R, Frazer CM, Grandi G;
XX MPI, 2000-6476q3/62.
XX Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections.
PS Claim 7; Appendix A; 692pp; English.
XX The present invention describes the full length genome of Neisseria
XX meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
XX represent fragments of the NMB genomic sequence, as the sequence was too
XX long to go in a record on its own it was split into 8 sequences which too
XX overlap each other at the beginning and end of each sequence by 49980 bp
XX (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
XX AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
XX AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
XX given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
XX primers which are used in the exemplification of the present invention.
XX The NMB genome and fragments from it have antibacterial activity, and can
XX be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
XX and/or antibodies which binds to the proteins can be used in compositions
XX for treating or preventing infection due to Neisserial bacteria or as a
XX diagnostic reagent for detecting the presence of Neisserial bacteria or
XX of antibodies raised to Neisserial bacteria. Computers, computer memory,
XX computer storage medium or computer databases can be used in a search to
XX identify open reading frames (ORFs) or coding sequences within the NMB
XX genome. The DNA sequences provide further opportunities to find antigenic
XX or immunogenic proteins which are more effective in vaccines than the
XX outer membrane proteins currently used
XX
SQ Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 U; 0 Other;
Query Match 99.8%; Score 826.4; DB 3; Length 349980;
Best Local Similarity 99.9%; Pred. No. 1.7e-240;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCAAAACACGTTATCAGCTTACGTTACGTCGCGAAGACGAGCGCGACATTGCCGAT 60
DB 226656 ATGCAAAACACGTTATCAGCTTACGTTACGTCGCGAAGACGAGCGCGACATTGCCGAT 226597
QY 61 ACCCTGCGGAGGACGAGCGATCCCGTTTTCAGTTTCGAGCGACGTATGCCGTGAAGG 120
DB 226596 ACCCTGCGGAGGACGAGCGATCCCGTTTTCAGTTTCGAGCGACGTATGCCGTGAAGG 226537
QY 121 CTGGAACAGGCAATGCGGCACTGTCGCCGCTGTGCGCGCACCCCTATTGAGCGGA 180
DB 226536 CTGGAACAGGCAATGCGGCACTGTCGCCGCTGTGCGCGCACCCCTATTGAGCGGA 226477
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCGGATTTGTGAAGACGAGCATTTGAGCAAGGT 240
DB 226476 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCGGATTTGTGAAGACGAGCATTTGAGCAAGGT 226417

QY 241 CTGCCGTATATACCGGTATTTGAGGACGAGCTTTTACTCGCGAAGATGAGGAAAAATTC 300
DB 226416 CTGCCGTATATACCGGTATTTGAGGACGAGCTTTTACTCGCGAAGATGAGGAAAAATTC 226357
QY 301 CTTCGCAAGACGCTTGCGCTGCAGACGCTTTGACCCCGATACCCGCTTTATCGTCGCG 360
DB 226356 CTTCGCAAGACGCTTGCGCTGCAGACGCTTTGACCCCGATACCCGCTTTATCGTCGCG 226297
QY 361 TTGGAAGAGATGTTATGACAGTCCTGACCTCGGCTCGGCGGAGATTAATGAGGAGG 420
DB 226296 TTGGAAGAGATGTTATGACAGTCCTGACCTCGGCTCGGCGGAGATTAATGAGGAGG 226237
QY 421 CGCGCTTTCCGCTGTTGAAAGCGAACACTGCGGGAAGCGGCGTATATGATTTCCCGA 480
DB 226236 CGCGCTTTCCGCTGTTGAAAGCGAACACTGCGGGAAGCGGCGTATATGATTTCCCGA 226177
QY 481 AAGCGATGCGGTTTTCCTGACAGGTTTCCGCTCGGCTCGGCGGAGGCTGACCCG 540
DB 226176 AAGCGATGCGGTTTTCCTGACAGGTTTCCGCTCGGCTCGGCGGAGGCTGACCCG 226117
QY 541 GTCGATCTGATGATGTTTCAGCGATTTTTCGACAGGAGAGATGCGGTTGCGAGCTC 600
DB 226116 GTCGATCTGATGATGTTTCAGCGATTTTTCGACAGGAGAGATGCGGTTGCGAGCTC 226057
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTTCACAGCAACAAACAGCGCA 660
DB 226056 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTTCACAGCAACAAACAGCGCA 225997
QY 661 TTGGCAGCTGATGCAACAGACCGCTCTCTGAAACGCAAAAGCGCGATTC 720
DB 225996 TTGGCAGCTGATGCAACAGACCGCTCTCTGAAACGCAAAAGCGCGATTC 225937
QY 721 CCCGCAACATTTCAACACCGCTGATCGCGCTTTGACCAATATCAGAGGAAAGG 780
DB 225936 CCCGCAACATTTCAACACCGCTGATCGCGCTTTGACCAATATCAGAGGAAAGG 225877
QY 781 GAAAAACGCGGCAAGGCGGCAACAGTTCATTGCTTCCTTCCAAATA 828
DB 225876 GAAAAACGCGGCAAGGCGGCAACAGTTCATTGCTTCCTTCCAAATA 225829
RESULT 5
AAT14061
ID AAT14061 standard; DNA; 5859 BP.
AC AAT14061;
XX 16-OCT-2003 (revised)
DT 09-JUN-1996 (first entry)
XX
XX N. gonorrhoeae 1gt locus.
XX
XX
XX Glycosyltransferase; 1,1po-oligosaccharide; 1gt gene; LOS locus; vaccine;
XX ds.
XX
XX Neisseria gonorrhoeae, strain F62.
OS
XX
XX
XX Key Location/Qualifiers
FH 445..1491
FT CDS
FT /*tag= a
FT /product= "LgtA"
FT 699..715
FT /*tag= b
FT /note= "poly-G tract"
FT 1491..2330
FT CDS
FT /*tag= c
FT /product= "LgtB"
FT 2342..3262
FT /*tag= d
FT /product= "LgtC"
FT 2499..2508
FT misc_feature
FT /*tag= e

```

FT      /note= "poly-G tract"
FT      3322..4335
FT      /*tag= f
FT      /product= "lgtD"
FT      misc_feature
FT      3576..3586
FT      /*tag= g
FT      /note= "poly-G tract"
FT      CDS
FT      4354..5196
FT      /*tag= h
FT      /product= "lgtE"
XX
XX      MO9610086-A1.
XX
XX      04-APR-1996.
XX
XX      25-SEP-1995; 95MO-US012317.
XX
XX      26-SEP-1994; 94US-00312387.
XX
XX      (UYRQ ) UNIV ROCKEFELLER.
XX
XX      Gotschlich EC;
XX
XX      MPI; 1996-200924/20.
XX      P-PSDB; AAR91311, AAR91312, AAR91313, AAR91314, AAR91315.
XX
XX      Nucleic acids encoding glycosyl transferase(s) - used in the diagnosis of
XX      infection with Neisseria and for the biosynthesis of oligo:saccharide(s).
XX
XX      Claim 1; Fig 29-m; 81pp; English.
XX
XX      The lgt locus (AAT14061) of Neisseria gonorrhoeae F62 contains 5 open
XX      reading frames, lgtA, lgtB, lgtC, lgtD and lgtE, coding for 5
XX      glycosyltransferases (see also AAR91311-15) involved in gonococcal
XX      lipooligosaccharide (LOS) biosynthesis. The sequence was constructed from
XX      2 clones isolated from an F62 gene bank in lambda-200 following screening
XX      with plasmid pRI0P1. 3 Of the coding sequences contain poly-G tracts that
XX      make them susceptible to premature termination. The lgt coding sequences
XX      can be used for prodn. of recombinant Lgt glycosyltransferases that are
XX      utilised in biosynthesis of LOS useful in vaccine prepn. (Updated on 16-
XX      OCT-2003 to standardise OS field)
XX
XX      Sequence 5859 BP; 1411 A; 1462 C; 1661 G; 1325 T; 0 U; 0 Other;
XX
XX      Query Match      87.7%; Score 726; DB 2; Length 5859;
XX      Best Local Similarity 93.2%; Pred. No. 1,4e-210;
XX      Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
XX
XX      1  ATSCAAACACGATGATACAGCTTAGCTTCGCGCGACGAAGCGAGGGGCGACATTGGCGGAT 60
XX      1491 ATGCAAAACCAAGCTTATACAGCTTGGCTTCGCGCGACGAAGCGAGGGGCGACATTGGCGGCA 1550
XX
XX      61  ACCCTTGGGACGAGCGGCGATCCCGTTTCAGTTTTCGACGCACTGATGCGTCTGAAGG 120
XX      1551 ACCTTGGGACGAGTGGCGGCACTCCGTTCCAGTTTTCGACGCACTGATGCGTCTGAAGG 1610
XX
XX      121  CTGGAACAGGCAATGGCGGCACTGTCGCCGCTTGTGGCGGCACTCCCTATTTGAGCGGA 180
XX      1611 CTGGAACGGGCAATGGCGGCACTGTCGCCGCTTGTGGCGGCACTCCCTATTTGAGCGGA 1670
XX
XX      181  GTGGAAGAAAGCGCTTTATGAGCGACGCGGTATTTGGAAGGAGGAGCATTTGAGCGAAGGT 240
XX      1671 GTGGAAGAAAGCGCTTTATGAGCGACGCGGTATTTGGAAGGAGGAGCATTTGAGCGAAGGT 1730
XX
XX      241  CTGCGCGTATATCAACGCTATTTGAGGACGATTTTATCTGCGCGAAGGTGAGAAAATTC 300
XX      1731 GTACCGTATATCGCCGATTTTGAAGATGATCTTACTGCGCGAAGGTGAGAAAATTC 1790
XX
XX      301  CTTCGCGGAAGAGCGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCGCCG 360
XX      1791 CTTCGCGGAAGATCTTCGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCGCCG 1850
XX
XX      361  TTGGAAGAGATGTTATGACGCTGACCTCGCCCTCGCGGCGTGGCGGATTAATCGCGG 420

```

```

Db      1851 TTGGAAGAGATGTTATGACGCTGACCTCGCCCTCGCGGCTGGCGGACTACGCGG 1910
XX
XX      421  CGGCGCTTTCCGCTGTGGAAAAGCAACCTGGGGGACCGCGGCTATATCATTTCCCGA 480
XX      1911 CGGCGCTTTCCGCTGTGGAAAAGCAACCTGGGGGACCGCGGCTATATCATTTCCCGA 1970
XX
XX      481  AAAGGAGAGCGGTTTTCTGACAGGTTTGGCGCCCTGGCCCGGAAAGGCTGACCCC 540
XX      1971 AAAGGAGAGCGGTTTTCTGACAGGTTTGGCGCCCTGGTTTGGCCGGAACGCTGACCC 2030
XX
XX      541  GTGATCTGATGATGTTTCAGCATTTTTCACAGGAGAGAAATGCGGTTTGCAGCTC 600
XX      2031 GTGATTTGATGATGTTTCAGCATTTTTCACAGGAGAGAAATGCGGTTTGCAGCTC 2090
XX
XX      601  AATCCCGCTTGTGGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 660
XX      2091 AATCCCGCTTGTGGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 2150
XX
XX      661  TTGGGAGAGCTGATGGAACAGACCGCCCTCTGAACCGGAAACAGCAAGGCGCATTC 720
XX      2151 TTGGGAGAGCTGATGGAACAGACCGCCCTCTGAACCGGAAACAGCAATGGCGCATTC 2210
XX
XX      721  CCGGCAACACATTCGAAACAGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
XX      2211 CCGGCAACACATTCGAAACAGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 2270
XX
XX      781  GAAAAACGCGGCAAGGCGGCAACAGTTCAATG 814
XX      2271 GAAAAACGCGGCAAGGCGGCAACAGTTAATCG 2304

```

```

XX      RESULT 6
XX      AAT49230
XX      ID AAT49230 standard; DNA; 5859 BP.
XX
XX      AAT49230;
XX
XX      AC 17-OCT-2003 (revised)
XX      DT 21-MAR-1997 (first entry)
XX
XX      DE lipo-oligosaccharide (including polyglycosyltransferase) gene.
XX
XX      KM Polyglycosyltransferase; N-acetylglucosaminyl transferase;
XX      KM N-acetylglucosaminyl transferase; lipo-oligosaccharide; ss.
XX
XX      OS Neisseria gonorrhoeae; ATCC 33084.
XX
XX      FH Key      Location/Qualifiers
XX      FT CDS      1..381
XX      FT          /*tag= a
XX      FT          445..1491
XX      FT          /*tag= b
XX      FT          /product= "polyglycosyltransferase"
XX      FT CDS      2342..3262
XX      FT          /*tag= c
XX      FT          3322..4335
XX      FT          /*tag= d
XX      FT          4354..5196
XX      FT CDS      /*tag= e
XX
XX      PN MO9640971-A1.
XX
XX      PD 19-DEC-1996.
XX
XX      PF 03-JUN-1996; 96MO-US008323.
XX
XX      PR 07-JUN-1995; 95US-00478140.
XX
XX      PA (NEOS-) NROSE TECHNOLOGIES INC.
XX
XX      PI Johnson KF, Roth S, Buczala SL;

```



```
QY 241 CTGCGTATATACCGTATTGAGACGACGTTTACTCGGCGAAGGTGAGAAAATTTC 300
DB 241 CTGCGGTATATCCCGGATTGAGACGACGTTTACTCGGCGAAGGTGAGACGACGTTTC 300
QY 301 CTGCGGAGACGCTTGCTGCAAGAACGCTTTGACCCGGATACCGCTTTATCTGTCGCG 360
DB 301 CTGCGGAGATATCTTGCTGCAAGACGCTTTGATTAAGATTTCCGCTTTATCTGTCGCG 360
QY 361 TTGGAAACGATGTTTATGACGCTCTGACCTCGGCTCGGCTGCGGATTAAGTGGG 420
DB 361 TTGGAAACGATGTTTATGACGCTCTGACCTCGGCTCGGCTGCGGATTAAGTGGG 420
QY 421 CGGCGCTTTCGCTGTTGAGAAAGGAACTGAGGAGGAGCGGCGGCTATATATTCCTCGA 480
DB 421 CGGCGCTTTCGCTGTTGAGAAAGGAACTGAGGAGGAGCGGCGGCTATATATTCCTCGA 480
QY 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGGCGGCGGAGGCTGCAACCC 540
DB 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGGCGGCGGAGGCTGCAACCC 540
QY 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGGAAATGCCGTTTCCAGCTC 600
DB 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGGAAATGCCGTTTCCAGCTC 600
QY 601 AATCCCGCTTGTGCGCGCAAGAGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
DB 601 AATCCCGCTTGTGCGCGCAAGAGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGGGAGCGCTGATTCGACACGACCGCTCTCTGACCGCAAGCAAGGAGGCGCATTC 720
DB 661 TTGGGAGCGCTGATTCGACACGACCGCTCTCTGACCGCAAGCAAGGAGGCGCATTC 720
QY 721 CCGCGCAACGATTCGAAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 721 CCGCGCAACGATTCGAAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
QY 781 GAAAAAGCGCGCAAGGCGCGCAAGCATTTG 814
DB 781 GAAAAAGCGCGCAAGGCGCGCAAGCATTCG 814

RESULT 8
ABZ38368
ID ABZ38368 standard; DNA; 840 BP.
AC ABZ38368;
AT 07-MAR-2003 (first entry)
CT N. gonorrhoeae nucleotide sequence SEQ ID 1325.
DE N. gonorrhoeae nucleotide sequence SEQ ID 1325.
XX Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX Neisseria gonorrhoeae.
OS Neisseria gonorrhoeae.
PN WO200279243-A2.
PD 10-OCT-2002.
PF 12-FEB-2002; 2002WO-IB002069.
PR 12-FEB-2001; 2001GB-00003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizsa M, Masignani V, Monaci E;
XX MPI; 2003-058415/05.
XX P-PSDB; ABP7398.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
```

```
XX PS Disclosure; page 284; 815bp; English.
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.
XX CC Also disclosed are the nucleic acid molecules encoding the proteins and
XX CC antibodies that specifically bind to the proteins. The composition
XX CC comprising the protein, nucleic acid or antibody is useful for the
XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae
XX CC infection, this may be in the form of a vaccine or gene therapy.
XX CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX CC molecules of the invention
SQ Sequence 840 BP; 209 A; 167 C; 242 G; 222 T; 0 U; 0 Other;
Query Match 57.0%; Score 471.8; DB 10; Length 840;
Best Local Similarity 80.7%; Pred. No. 2.7e-133;
Matches 551; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 1 ATGCAAAACACGATTATGAGCTTAGCTTCGCGCGAAGACGAGGCGGCAATGCGCAT 60
DB 1 ATGCAAAACACGATTATGAGCTTAGCTTCGCGCGAAGACGAGGCGGCAATGCGCAT 60
QY 61 ACCCTCGGAGGACGCGCATCCGCTTTCAGTTTTCAGCACTGATGCCGCTGCAAGG 120
DB 61 ACCCTCGGAGGACGCGCATCCGCTTTCAGTTTTCAGCACTGATGCCGCTGCAAGG 120
QY 121 CTGGAACAGGCAATGCGGGAATCGTCCCGGCTGTGCGGCAACCCCTATTGAGCGGA 180
DB 121 CTGGAACAGGCAATGCGGGAATCGTCCCGGCTGTGCGGCAACCCCTATTGAGCGGA 180
QY 181 GTGAAAAAGCGCTGTTATGAGCAAGCGGATTTGTAAGCAAGCAATGGAAGAAAGT 240
DB 181 GTGAAAAAGCGCTGTTATGAGCAAGCGGATTTGTAAGCAAGCAATGGAAGAAAGT 240
QY 241 CTGCGTATATACCGTATTTGAGACGACGTTTACTCGGGAAGGATGAGAAATTC 300
DB 241 CTGCGTATATACCGTATTTGAGACGACGTTTACTCGGGAAGGATGAGAAATTC 300
QY 301 CTGCGCAAGACGCTTGCTGCAAGACGCTTGAACCGGATACCGGCTTTATCGTCGCG 360
DB 301 CTGCGCAAGACGCTTGCTGCAAGACGCTTGAACCGGATACCGGCTTTATCGTCGCG 360
QY 361 TTGGAAGAGATGTTATGACAGCTCCGACCTCGGCTCGGCGGATTAAGTAAAG 420
DB 361 TTGGAAGAGATGTTATGACAGCTCCGACCTCGGCTCGGCGGATTAAGTAAAG 420
QY 421 CGGCGCTTTCGCTGTTGGAAGGCAAGCACTGAGGAGCGGCTATATTCGCGA 480
DB 421 CGGCGCTTTCGCTGTTGGAAGGCAAGCACTGAGGAGCGGCTATATTCGCGA 480
QY 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGGCGGCAAGGCTGCAACCC 540
DB 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGGCGGCAAGGCTGCAACCC 540
QY 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGGAAATGCCGTTTCCAGCTC 600
DB 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGGAAATGCCGTTTCCAGCTC 600
QY 601 AATCCCGCTTGTGCGCGCAAGAGCTGCAATTAAGCAAGTTTTCAGCAAGCAAAACAGCGCA 660
DB 601 AATCCCGCTTGTGCGCGCAAGAGCTGCAATTAAGCAAGTTTTCAGCAAGCAAAACAGCGCA 660
QY 661 TTGGGAGCGCTGATTCGAAACAGCA 683
DB 661 TTGGGAGCGCTGATTCGAAAGAAAGA 683

RESULT 9
ABX09919
ID ABX09919 standard; DNA; 840 BP.
XX ABX09919;
XX AC
```

DT 27-OCT-2003 (revised)
 DT 22-JAN-2003 (first entry)
 XX N. meningitidis DNA encoding a vaccine antigen #61.
 DE Vaccine; antigen; ds; meningococcal disease; pathogenic bacteria;
 KM meningitis.
 XX
 OS Neisseria meningitidis serogroup B.
 PN W020027648-A2.
 PD 03-OCT-2002.
 XX
 PF 22-MAR-2002; 2002WO-GB001399.
 XX
 PR 22-MAR-2001; 2001GB-00007219.
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PI Robinson A, Goringe AR, Hudson MJ, Bracegirdle P, West DM;
 PI Oliver KJ, Kroll JS, Langford PR;
 XX
 DR WPI; 2003-018958/01.
 DR P-PSDB; AB006084.
 XX
 PT Identifying an antigen for manufacturing a vaccine against meningococcal
 PT infection, comprises contacting antibodies with polypeptides, detecting
 PT polypeptide-antibody complexes, and identifying bound polypeptides as
 PT antigens.
 XX
 PS Claim 34; Page 278-279; 310pp; English.
 CC The invention relates to identifying an antigen comprising: (a) obtaining
 CC antibodies against a commensal bacteria, or an extract from a commensal
 CC bacteria; (b) contacting the antibodies with polypeptides obtained from
 CC an expression library of either a commensal or a pathogenic bacteria; (c)
 CC determining whether the polypeptides bind to antibodies; and (d) (where a
 CC polypeptide binds to an antibody) identifying that polypeptide as an
 CC antigen. Also included are: (1) a method of preparing a vaccine
 CC composition, comprising identifying an antigen with the above method, and
 CC combining the antigen with a carrier; (2) a vaccine composition obtained
 CC by the above methods; (3) an antigen identified by the above method; (4)
 CC a polypeptide encoded by all or a part of a nucleic acid sequence
 CC comprising the Neisseria lactamica DNA sequences detailed in the
 CC specification; (5) a vector comprising the nucleic acid molecule; (6) a
 CC method of preparing a composition for vaccination against infection by
 CC pathogenic bacteria, comprising: (a) obtaining a first antigen from a
 CC commensal Neisseria; (b) comparing the amino acid sequence of the first
 CC antigen with the amino acid sequence of the second antigen from a
 CC pathogenic bacteria, or comparing the sequence of a nucleic acid which
 CC codes for the first antigen with the sequence of the nucleic acid that
 CC codes for the second antigen; and if the first antigen is homologous to
 CC the second antigen or if the nucleic acid sequence for the first antigen
 CC is homologous to that of the second antigen, and (c) preparing a
 CC composition for vaccination against bacterial infection comprising the
 CC first antigen; and (7) an antibody that binds to the polypeptide antigen.
 CC The method is useful in screening commensal and pathogenic bacteria for
 CC previously unidentified vaccine antigens by identifying polypeptide
 CC antigens that bind to sera raised against commensal bacterial proteins.
 CC The polypeptide is useful as a vaccine antigen which may be used in the
 CC manufacture of a medicament for vaccination against meningococcal
 CC infection (e.g. meningitis). The present sequence encodes an antigenic
 CC protein from the pathogenic bacteria N. meningitidis. (Updated on 27-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 840 BP; 218 A; 154 C; 244 G; 224 T; 0 U; 0 Other;
 Query Match 51.4%; Score 426; DB 8; Length 840;
 Best Local Similarity 75.2%; Pred. No. 2.7e-119;
 Matches 531; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 1 ATCAAAAACGATTATGAGTTAGCTTCCGCGGAGAACGAGGAGCGACATTGCCAT 60
 DB 1 ATCAAAAACGATTATGAGTTAGCTTCCGCGGAGAACGAGGAGCGACATTGCCAT 60
 QY 61 ACCCTGGGAGGACGCGCATCCGTTTCAAGTTTTCAGACGCACTGATCCGCTGAAAG 120
 DB 61 ACCCTGGGAGGAGGAGCATCCGTTTCAAGTTTTCAGACGCACTGATCCGCTGAAAG 120
 QY 121 CTGGAACGAGCAATGGCGGAAGCTGTCGCCGCTGTGGGCGACCCCTATTGAGCGGA 180
 DB 121 CTGAATCGGATGATGGCGGAATGGTGGCGGATTTGGCAAAACAGACCTGTGGAGCA 180
 QY 181 GTGAAAAAAGCTGCTTTATGAGCCAGCCGATTTGTGAAAGCAGCATTGACCAAGAT 240
 DB 181 GTGAAAAAAGCTGCTTTATGAGCCAGCCGATTTGTGAAAGCAGCATTGATGAAGT 240
 QY 241 CTGCGGATATATCCGTTATTTGAGAGCAGCTTTTACTCGCGAAGGTGAGAAAAATTC 300
 DB 241 CTGCGCTATGTTGCGGATTTGAGAGATGCTCTGCTTGCAGAAAGCAGAAAAAGTTC 300
 QY 301 CTGCGCAAGACGCTTGGCTGCAAGACGCTTTCACCCGATACCGCTTATCGTCCGC 360
 DB 301 CTGCGCAAGATCTTGGCTTGAAGAGCGCTTGTGATAGATTTGCTTATCGTTCGT 360
 QY 361 TTGGAACGATGTTATGCAAGTCTGACCTTGCCCTCCGCGTGGCGGATTAATGCGAG 420
 DB 361 TTGGAACGATGTTATGCAAGTATGTCAGACCGGATTAAGTCTGAATTAATGAGAT 420
 QY 421 GCGGCTTTCGCTTGTGAAAGCAACGCGGGGAGCGGGGTATATGATTTCCCGCA 480
 DB 421 GCGGCTTTCGCTTGTGAAAGCAACGCGGGGAGCGGGGTATATGATTTCCCGCA 480
 QY 481 AAAGGATGCGGTTTCTGAGACAGTTTCCGCGCTGCGCGCGAAGGCGCTGACCCC 540
 DB 481 GAGGCGATGCGGTTTCTTGGAAAGGTTTCCGCTTTCGCGGAGATGATTAAGCG 540
 QY 541 GTGCAATGATGATGTTGAGGATTTTTCAGACGAGGAAGAAATCCGTTTGCAGTCC 600
 DB 541 GTGATGATGATGTTTACTTATTTCTTATTAAGGAGGAGATGCTGTTATCAAGTT 600
 QY 601 AATCCGCGCTTGTGCGCCAGAGAGCTGATATCCCAAGTTTCAAGCAACCAAGCGCA 660
 DB 601 AATCCGCGCTTATGATACCAAGAAATGATATATGCAAGTTCTTCAGTAAAAAGTATG 660
 QY 661 TTGGGACGCTGATGAAACAGACCGCTCTGAAACGCAACGCC 706
 DB 661 TTGGGTAGCGATTGGAAGAAATAGGAAACAGAAAGACACC 706
 RESULT 10
 ID ADT05534 standard; DNA; 14547 BP.
 XX ADT05534;
 AC ADT05534;
 XX 02-DEC-2004 (first entry)
 DT Haemophilus influenzae (NHI) contig DNA sequence - SEQ ID 570.
 DE Haemophilus influenzae (NHI) contig DNA sequence - SEQ ID 570.
 KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
 KM contig.
 XX
 OS Haemophilus influenzae.
 XX
 PN W02004078949-A2.
 PD 16-SEP-2004.
 XX
 PF 05-MAR-2004; 2004WO-US007001.
 XX
 PR 06-MAR-2003; 2003US-0453134P.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL INC.

XX Bakaletz LO, Munson RS, Dyer DW;
XX WPI; 2004-662422/64.
XX
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
XX Example 1; SEQ ID NO 570; 88bp; English.
XX
XX The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi config sequence of the invention.
XX
XX Sequence 14547 BP; 4491 A; 2521 C; 2901 G; 4631 T; 0 U; 3 Other;
SQ
Query Match 11.2%; Score 93; DB 13; Length 14547;
Best Local Similarity 51.9%; Pred. No. 4.7e-17;
Matches 235; Conservative 0; Mismatches 215; Indels 3; Gaps 1;
QY 180 AGTGAAGAAAGCCTGCTTTATGAGCCAGCCGATTTGTGAAAGCAGCATTTGACGAAAG 239
DB 11201 AGGAGAAAAGGCTGTTTATTAAGCCATTCTTATTAAGAAATAGTGTGAATGAAAA 11260
QY 240 TCTGCGATATATCAACCGTATTTGAGACGAGCTTTTACTCGGCGAAGGTGAGAAAAATT 239
DB 11261 TTGGATATATCTCAATTTTGAAGATGATGATATTTGGCGAAGATGCGGAAGTGT 11320
QY 300 CATTGCCGAAGAGCGTTGGCTGCAAGAACGTTGACCCGAGTACCCCTTTATCGCCG 359
DB 11321 TTGGCAACAAGATGAATGGTTAAACACGTTTGATTTTAATGAATATTTTATATTG 11380
QY 360 CTGGAACAGATGTTATGACAGT---CTGACCTCGCCCTCGGCGTGGCGGATTAAGT 416
DB 11381 TTGAAGAACTTTTAAAGCAAGTTAACTTGAAGAAACAATTAATTCACCTTTTA 11440
QY 417 CGGCGCGGCTTTCCGCTGTGGAAGCAACTGGGGGACGCGCGGCTATATCATTTTC 476
DB 11441 TTCTAGGAACCTTTGATATTTTAAATGACCTACCTGGGGGACGCGGTATATATTTC 11500
QY 477 CCGAAAAGCGATGCGGTTTTCCTGACAGTTTCCGCGCTGCGCGCGAAGGCTGCA 536
DB 11501 TCAAGGTGCGGCTAAATATGTAATTAATTAATTAATTAATTAATTAATTAAT 11560
QY 537 CCGCGTCAATCTGATGATGATGTTCAAGCAATTTTTCGACAGGAAAGAAATGCGGTTGCCA 596
DB 11561 TGCAGTTGATGAATTTATTTTAAATTAATTAATTAATTAATTAATTAATTAAT 11620
QY 597 GCTCAATCCGCGCTTGTGCGCCCAAGAGCTGCA 629
DB 11621 ACTTAATCCAGCAATTTGATTCAGAACTCCA 11653
RESULT 11
ADT05645
ID ADT05645 standard; DNA; 106645 BP.
XX
XX ADT05645;
AC
XX
XX 02-DEC-2004 (first entry)
XX
XX Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 681.
DE
XX
XX middle ear bacterial infection; nasopharynx bacterial infection; ds.
KM
XX
XX Haemophilus influenzae.
OS
XX
XX WO2004078949-A2.
PN
XX
XX 16-SEP-2004.
PD

XX
PF 05-MAR-2004; 2004MO-US007001.
XX
XX 06-MAR-2003; 2003US-0453134P.
XX
XX (CHIL-) CHILDRENS HOSPITAL INC.
XX
XX Bakaletz LO, Munson RS, Dyer DW;
XX
XX WPI; 2004-662422/64.
XX
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
XX Claim 1; SEQ ID NO 681; 88bp; English.
XX
XX The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi DNA sequence of the invention.
XX
XX Sequence 106645 BP; 3261 A; 1943 C; 21227 G; 33322 T; 0 U; 0 Other;
SQ
Query Match 11.2%; Score 93; DB 13; Length 106645;
Best Local Similarity 51.9%; Pred. No. 1.1e-16;
Matches 235; Conservative 0; Mismatches 215; Indels 3; Gaps 1;
QY 180 AGTGAAGAAAGCCTGCTTTATGAGCCAGCCGATTTGTGAAAGCAGCATTTGACGAAAG 239
DB 98869 AGGAGAAAAGGCTGTTTATTAAGCCATTCTTATTAAGAAATAGTGTGAATGAAAA 98928
QY 240 TCTGCGATATATCAACCGTATTTGAGACGAGCTTTTACTCGGCGAAGGTGAGAAAAATT 239
DB 98929 TTGGATATATCTCAATTTTGAAGATGATGATATTTCTGGCGAAGATGCGGAAGTGT 98988
QY 300 CATTGCCGAAGAGCGTTGGCTGCAAGAACGTTGACCCGAGTACCCCTTTATCGCCG 359
DB 98989 TTGGCAACAAGATGAATGGTTAAACACGTTTGATTTTAATGAATATTTTATATTG 99048
QY 360 CTGGAACAGATGTTATGACAGT---CTGACCTCGCCCTCGGCGTGGCGGATTAAGT 416
DB 99049 TTGAAGAACTTTTAAAGCAAGTTTAACTTGAAGAAACAATTAATTCACCTTTTA 99108
QY 417 CGGCGCGGCTTTCCGCTGTGGAAGCAACTGGGGGACGCGCGGCTATATCATTTTC 476
DB 99109 TTCTAGGAACCTTTGATATTTTAAATGACCTACCTGGGGGACGCGGTATATATTTC 99168
QY 477 CCGAAAAGCGATGCGGTTTTCCTGACAGTTTCCGCGCTGCGCGCGAAGGCTGCA 536
DB 99169 TCAAGGTGCGGCTAAATATGTAATTAATTAATTAATTAATTAATTAATTAAT 99228
QY 537 CCGCGTCAATCTGATGATGATGTTCAAGCAATTTTTCGACAGGAAAGAAATGCGGTTGCCA 596
DB 99229 TGCAGTTGATGAATTTATTTTAAATTAATTAATTAATTAATTAATTAATTAAT 99288
QY 597 GCTCAATCCGCGCTTGTGCGCCCAAGAGCTGCA 629
DB 99289 ACTTAATCCAGCAATTTGATTCAGAACTCCA 99321
RESULT 12
AAT42063_05/c
Continuation (6 of 19) of AAT42063 from base 500001 (Haemophilus influenzae complete ge
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063
WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000

WP AAT42063_06 600001 710000
 WP AAT42063_07 700001 810000
 WP AAT42063_08 800001 910000
 WP AAT42063_09 900001 1010000
 WP AAT42063_10 1000001 1110000
 WP AAT42063_11 1100001 1210000
 WP AAT42063_12 1200001 1310000
 WP AAT42063_13 1300001 1410000
 WP AAT42063_14 1400001 1510000
 WP AAT42063_15 1500001 1610000
 WP AAT42063_16 1600001 1710000
 WP AAT42063_17 1700001 1810000
 WP AAT42063_18 1800001 1830121

Query Match 10.5%; Score 86.6; DB 2; Length 110000;
 Best Local Similarity 51.0%; Pred. No. 1e-14; Mismatches 219; Indels 3; Gaps 1;
 Matches 221; Conservative 0;
 QY 180 AGTGAAGAAAGCTGCTTTATGAGCCAGCCGCTATTGTGGAAGCAGGCAATTGACGAGG 239
 Db 70747 AGGAGAAAGAGCTGTTTATGAGCCATTTCTATATGGAATTAAGTGTGATGAA 70688
 QY 240 TCTGCCGTATATCAGCGTATTTGAGGAGCAGACGCTTTTACTCGCGGAAGTGAGGAAATTT 299
 Db 70687 TTTGAAATATCTCAAAATTTTGAAGATGATGTAATTTCTTGCGGAATGCGGAATATT 70628
 QY 300 CCTTGCAGAGAGCTTGTGCTGCAAGAGCGCTTTGACCCGSAATACCGCTTTATCGTCCG 359
 Db 70627 TTTGAACCAAAATGATGTTTAAACACGTTTGTATTTTANATTTTATTTATTTG 70568
 QY 360 CTTTGGAAGAGATGTTATGACGT--CCTGACCTGCGCTCGCGCTGCGGATTTACTG 416
 Db 70567 TTTGAAACCTTTTTCACCCAGCTTAACTTTGAGAAACAACTTAAATTCACCTTTTAA 70508
 QY 417 CGGCGCGGCTTTCGCTGTGGAAGCAGACACTGCGGAGCGCGGCTATATCATTTTC 476
 Db 70507 TTTCTAGGAATTTTGAATTTTAAATCACTGCGGAGCGGAGTTATTTATTTTC 70448
 QY 477 CCGAAAGCGATGCGGCTTTTCTGTGACAGGTTTGCCTGCGCCCGGAAAGGCTGCA 536
 Db 70447 TCAGAGTGGGCTTAATATGTAATGAAATTAATTAATTAATTAATTAATTAATTTGT 70388
 QY 537 CCCCGCATCTGATGATGTTGCGGATTTTTCGACAGGAAAGAAATCCCGTTTGCCA 596
 Db 70387 TGCAGTTGATGAATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTTGCTATCA 70328
 QY 597 GCTCAATCCGCTTGTGCGCCCAAGAGCTGCA 629
 Db 70327 ACTGAATCCAGCAATTTGTATTCAGAACTCCA 70295

RESULT 13
 ADL02801
 ID ADL02801 standard; DNA; 768 BP.
 AC ADL02801;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE DNA encoding a M. catarrhalis protein #487.
 XX
 KW ds; gene; Moraxella catarrhalis; infection.
 XX
 OS Moraxella catarrhalis.
 XX
 PN US6673910-B1.
 XX
 PD 06-JAN-2004.
 XX
 PF 04-APR-2000; 2000US-00540236.
 XX
 PR 08-APR-1999; 99US-0128416P.
 XX

PA (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton GL;
 PI
 XX WPI; 2004-178127/17.
 DR P-PSDB; ADL04721.
 XX
 PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
 PT preparing a composition for diagnosing, preventing or treating infection
 PT caused by Moraxella catarrhalis.
 XX
 PS Disclosure; SEQ ID NO 487; 429bp; English.
 XX
 CC The invention relates to an isolated nucleic acid encoding a Moraxella
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
 CC composition for diagnosing, preventing or treating infection caused by
 CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
 CC catarrhalis protein.
 XX
 SQ Sequence 768 BP; 270 A; 142 C; 131 G; 225 T; 0 U; 0 Other;

Query Match 8.8%; Score 73.2; DB 12; Length 768;
 Best Local Similarity 52.9%; Pred. No. 1.5e-11;
 Matches 181; Conservative 0; Mismatches 158; Indels 3; Gaps 1;
 QY 1 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAGAACGCGGCGCACATTGCCGAT 60
 Db 7 ATACAAATTTTGTATGATGATGTAATAAATCTGCCAATAAAGAGAACATTTATGTGT 66
 QY 61 ACCTTGGCAGGACGAGGATCCGTTTCACTTTTTCAGCACTGATGCGCTGGAAGG 120
 Db 67 GAATTTGGCAACAGGAGGATGCTTTGAGTGTGAGTAAACCCACCTGATTTT 126
 QY 121 CTGCAAGAGCAATGCGGAATCGTCCCGCTTGTGCGGCAACCCCTATTGAGCGGA 180
 Db 127 AGCAAGTACGCCCAAAAGCTTTCAATCCATCAACA---ACCAAGACTCACCGAT 183
 QY 181 GTGAAAAAGCTGCTTTATGAGCCAGCGCTATTGTGGAAGAGGCAATTGACGAGGT 240
 Db 184 GCGCAAAAAGCTGCTTTTAAAGCATGATGACATGAGCAACAGATGATGATGAAAAAC 243
 QY 241 CTGCGCTATATCACCGTATTTAGAGAGAGCTTTTACTCGCGGAGGTGAGGAAAAATTC 300
 Db 244 TTGATTTATGAGCAATTTTGAAGAGATGTCATTTTGGGAAATGACTTCAACAATTC 303
 QY 301 CTTGCCGAAGACGCTTGTGCAAGAACGCTTTGACCCGAT 342
 Db 304 TTACAAAGAACTGACAAATTTGGCTACAAACAATGATGCGTTGAT 345

RESULT 14
 ADV09334/c
 ID ADV09334 standard; DNA; 4583 BP.
 AC ADV09334;
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE M. catarrhalis strain 7169 serotype B LOS gene cluster.
 XX
 KW ds; diagnosis; bacterial infection; otitis media; LOS gene cluster.
 XX
 OS Moraxella catarrhalis.
 XX
 PN W02004I04170-A2.
 XX
 PD 02-DEC-2004.
 XX
 PF 13-MAY-2004; 2004WO-US015141.
 XX
 PR 13-MAY-2003; 2003US-0470022P.
 XX
 PA (UNIV) UNIV NEW YORK STATE RES FOUND.

Db 2791 TTGGATTATATGGCAATTTTGAAGACGATGTCTATTGGGAAATGACTCTCAACAATTC 2732
Qy 301 CTTGCCGAAGACGCTTGGCTGCAAGAACGCTTTGAACCCGAT 342
Db 2731 TTACAAGAACTGACAATTTGGCTACACAATAATGCCGTTGAT 2690

Search completed: April 7, 2006, 16:17:30
Job time : 938.991 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:00:12 ; Search time 6502.61 seconds
(without alignments)
5957.564 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacacagctatcacg.....tcattgtgccttccaataa 828

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	40.2	4.9	522	6	CD231899 SS1_30_A0
C 2	40.2	4.9	598	6	CD231864 SS1_30_A0
C 3	40.2	4.9	620	6	CB926280 ABA1_7_H0
C 4	40.2	4.9	729	6	CP430189 PRT_36_H0
C 5	39.2	4.8	1217	8	DR135967 43289438
C 6	39.2	4.7	530	1	AA461978
C 7	39.2	4.7	1357	9	CC306269 CH261-2K1
C 8	39.2	4.7	737	9	CC821108 FPCP4G8 U
C 9	38.6	4.7	582	3	BP210238
C 10	38.6	4.7	469	1	AM677069
C 11	38.6	4.7	775	9	BZ577887 msh2_5621
C 12	38.4	4.6	920	4	CNS0F4H1
C 13	38.2	4.6	569	7	CNS06175
C 14	38.2	4.6	593	6	CF687694
C 15	38.2	4.6	591	6	CF712358
C 16	38.2	4.6	597	6	CF693664
C 17	38.2	4.6	619	8	DR826908
C 18	38.2	4.6	666	6	CF678610
C 19	38.2	4.6	693	6	CF709985
C 20	38.2	4.6	715	6	CF695835
C 21	38.2	4.6	735	6	CF679141
C 22	38.2	4.6	743	6	CF713838

23	38	4.6	743	6	CF721366	CF721366 CCAAX90TR
24	38	4.6	749	6	CF714073	CF714073 CCAAB78TR
25	38	4.6	754	8	DR954928	DR954928 ZM BRB004
26	38	4.6	755	6	CF696727	CF696727 CCAHP79TR
27	38	4.6	757	6	CF675743	CF675743 CCAD077TR
28	38	4.6	760	9	BZ538612	BZ538612 OGAI587TC
29	38	4.6	762	6	CF678653	CF678653 CCAHV89TR
30	38	4.6	763	7	CO526725	CO526725 3530_1_17
31	38	4.6	779	6	CF701635	CF701635 CCA5322TR
32	38	4.6	781	6	CF676109	CF676109 CCAHV75TR
33	38	4.6	782	6	CF688826	CF688826 CCAD047TR
34	38	4.6	789	6	CF690987	CF690987 CCAHX83TR
35	38	4.6	793	6	CF676377	CF676377 CCAHP89TR
36	38	4.6	812	6	CF719729	CF719729 CCAK73TR
37	38	4.6	816	6	CF682178	CF682178 CCAI466TR
38	38	4.6	827	6	CF676273	CF676273 CCAD013TR
39	38	4.6	829	6	CF712852	CF712852 CCAH251TR
40	38	4.6	829	6	CF718427	CF718427 CCAH177TR
41	38	4.6	859	6	CF698378	CF698378 CCAH057TR
42	38	4.6	862	6	CF696539	CF696539 CCAHV93TR
43	38	4.6	863	6	CF696079	CF696079 CCAHB95TR
44	38	4.6	870	6	CF696612	CF696612 CCAD102TR
45	38	4.6	872	6	CF683469	CF683469 CCAH55TR

ALIGNMENTS

RESULT 1
CD231899/c
LOCUS
DEFINITION
SS1_30_A04.g1 A012 Salt-stressed seedlings Sorghum bicolor cDNA
Clone SS1_30_A04_A012 5', mRNA sequence.
ACCESSION
CD231899
VERSION
CD231899.1 GI:30975364
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Bukarjota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 522)
REFERENCE
Cordonnier-Pratt M.-M., Wentzel V., Suzuki Y., Sugano S.,
Klein R.R., Liang C., Sun F., Sullivan R., Shah M., Sumner E.J.,
Bastman A. and Pratt L.H.
An EST database from Sorghum: salt-stressed seedlings
Unpublished (2003)
TITLE
JOURNAL
COMMENT
Other ESTs: SS1_30_A04.b1 A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpat@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAGAGCTGCG).
Location/Qualifiers
1..522
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cui_type="IS3620C"
/db_xref="taxon:4558"
/clone="SS1_30_A04_A012"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Salt-stressed seedlings"

FEATURES

source

/note="Vector: PMB18s-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from 9-day-old seedlings grown in hydroponic culture. Seedlings were transferred to a 150 mM NaCl solution and harvested at 3, 6, 12 and 24 hr following transfer. Roots and leaves were pooled from all time points and RNA isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the PMB18s-FL3 vector (5'-prime DraIII site is CACTGATG, 3'-prime DraIII site is CACCATG)."

ORIGIN

Query Match 4.9%; Score 40.2; DB 6; Length 522;
Best Local Similarity 49.3%; Pred. No. 2.2;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGAGAGGAGGATTTGACGAGAGTCTCCGTATATACCGTATTTGAGGACGCTTT 274
DB 415 TCTGAGAGCTGTTTGGGAGAGGAGAACTGATGTTCCAAATGTTCTGAGGACCACTAACT 356
QY 275 TACTCGGCGAAGGTGAGGAAAAATTCCTTGCCGAGAGACGCTTGCAAGACGCTTTG 334
DB 355 TGTCTCGAGAGCGGCTCTTACCAAGCTTTGCCGATGTCCTTGAGCCGAGAACTCCTCTT 296
QY 335 ACCCGAATACCGCTTTATGCTTCGCTTGCGGAGAAAGATTTATGCACTGCTGACCTGCG 394
DB 295 AGCAGCAGAGTGTCTGCTTGTGAGCGGAGTGGCGGAGTCTGCGCGCTGAGGCGCCGA 236
QY 395 CCTCGCGGTGCGGATTAATGCGGCGCGCT 427
DB 235 GCGGCGGAGGAGGAGTGGCGGTGAGGAGCACT 203

RESULT 2
LOCUS CD231864/c 598 bp mRNA linear EST 21-MAY-2003
DEFINITION SSI_30_A08_g1_A012 Salt-stressed seedlings Sorghum bicolor CDNA
CD231864
VERSION CD231864.1 GI:30975329
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 598)
Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Summer, E.J., Eastman, A. and Pratt, L.H.
An EST database from Sorghum: salt-stressed seedlings
Unpublished (2003)

TITLE
JOURNAL
COMMENT
Other_ESTs: SSI_30_A08_b1_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

FEATURES
source
Email: mmpat@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).
Location/Qualifiers
1..598
/organism="Sorghum bicolor"
/mol_type="mRNA"

/cultivar="HS3620C"
/db_xref="taxon:4558"
/clone="SSI_30_A08_A012"
/lab_host="DH10B-rT phage-resistant E. coli"
/clone_id="Salt-stressed seedlings"
/note="Vector: PMB18s-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from 9-day-old seedlings grown in hydroponic culture. Seedlings were transferred to a 150 mM NaCl solution and harvested at 3, 6, 12 and 24 hr following transfer. Roots and leaves were pooled from all time points and RNA isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the PMB18s-FL3 vector (5'-prime DraIII site is CACTGATG, 3'-prime DraIII site is CACCATG)."

ORIGIN

Query Match 4.9%; Score 40.2; DB 6; Length 598;
Best Local Similarity 49.3%; Pred. No. 2.3;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGAGAGGAGGATTTGACGAGAGTCTCCGTATATACCGTATTTGAGGACGCTTT 274
DB 426 TCTGAGAGCTGTTTGGGAGAGGAGAACTGATGTTCCAAATGTTCTGAGGACCACTAACT 367
QY 275 TACTCGGCGAAGGTGAGGAAAAATTCCTTGCCGAGAGACGCTTGCAAGACGCTTTG 334
DB 366 TGTCTCGAGAGCGGCTCTTACCAAGCTTTGCCGATGTCCTTGAGCCGAGAACTCCTCTT 307
QY 335 ACCCGAATACCGCTTTATGCTTCGCTTGCGGAGAAAGATTTATGCACTGCTGACCTGCG 394
DB 306 AGCAGCAGAGTGTCTGCTTGTGAGCGGAGTGGCGGAGTCTGCGCGCTGAGGCGCCGA 247
QY 395 CCTCGCGGTGCGGATTAATGCGGCGCGCT 427
DB 246 GCGGCGGAGGAGGAGTGGCGGTGAGGAGCACT 214

RESULT 3
LOCUS CB926280/c 620 bp mRNA linear EST 28-APR-2003
DEFINITION ABA1_7_H09_g1_A012 Abscisic acid-treated seedlings Sorghum bicolor CDNA clone ABA1_7_H09_A012 5', mRNA sequence.
CB926280
VERSION CB926280.1 GI:30162551
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 620)
Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Buchanan, C.D., Eastman, A. and Pratt, L.H.
An EST database from Sorghum: ABA1-treated seedlings
Unpublished (2003)

TITLE
JOURNAL
COMMENT
Other_ESTs: ABA1_7_H09_b1_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

FEATURES
source
Email: mmpat@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

FEATURES
Seq primer: Sug5 (CTTCTGCTCTAAGAGTCGC).
Location/Qualifiers

source

1. .620
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone="ABAI 7 H09 A012"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The library was prepared from polyA+ RNA from seedlings grown in hydroponic culture. After 12 days, medium was supplemented with 1 mM abscisic acid (ABA), while leaves were misted with a solution of 1 mM ABA. Roots and leaves were harvested after 3, 6, 12, and 24 hr and material from all time points was combined prior to RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 4.9%; Score 40.2; DB 6; Length 620;
Best Local Similarity 49.3%; Pred. No. 2.3;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGGAAGCAGCATTTGAGAGAGGTGCGCTATACCGTATTGAGAGACGCTT 274
DB 446 TCTGGAAGCTGTTTGGGAGACGAGAACTGATGTTCCATGTTCTGAGACCACTA 387
QY 275 TACTCGGCGAAGGTGAGAAAATTCCTGCGGAGACGCTTGCTGCAAGACGCTTG 334
DB 386 TGTCTCGAAGCGGTCTCTAACAAGCTTCCGATGCTGTGAGCCAGAAATCGCTCT 327
QY 335 ACCCGATACCGCTTATGCTCGCTTGAGAAAGATGTTATGACAGTCTGACCTGCG 394
DB 326 AGCAGCAGAGTGTCTGCTGTGAGCGGCTGCGGAGATGCTCGGCTGAGAGCCCGCA 267
QY 395 CTTCCGCGCTGCGGATTAATCTGCGGCGCGCT 427
DB 266 GCGGCGCAGGAGGTGCGGCTGAGAGCAGCT 234

RESULT 4
LOCUS CP430189/c 729 bp mRNA linear EST 03-SEP-2003
DEFINITION PH1_26_H07_g1_A002 Phosphorous-deficient seedlings Sorghum bicolor
ACCESSION CP430189
VERSION CP430189.1 GI:34442890
KEYWORDS EST.

SOURCE
ORGANISM Sorghum bicolor (sorghum)
Sorghum bicolor
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
AUTHORS Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R. R., Liang, C.,
Sun, F., Sullivan, R., Harris, K., Eastman, A. and Pratt, L. H.
An EST database from Sorghum: phosphorous-deficient seedlings
Unpublished (2003)
COMMENT Other_ESTs: PH1_26_H07_b1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;

sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAGAGTCGC).
Location/Qualifiers

FEATURES

source

1. .729
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone="PH1_26_H07 A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The library was prepared from polyA+ RNA from Brx623 seedlings grown hydroponically in the absence of added phosphorous. At 14 days of age, roots and shoots were harvested and stored at -80 C until RNA was isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 4.9%; Score 40.2; DB 6; Length 729;
Best Local Similarity 49.3%; Pred. No. 2.4;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGGAAGCAGCATTTGAGAGAGGTGCGCTATACCGTATTGAGAGACGCTT 274
DB 431 TCTGGAAGCTGTTTGGGAGACGAGAACTGATGTTCCATGTTCTGAGACCACTA 372
QY 275 TACTCGGCGAAGGTGAGAAAATTCCTGCGGAGACGCTTGCTGCAAGACGCTTG 334
DB 371 TGTCTCGAAGCGGTCTCTAACAAGCTTCCGATGCTGTGAGCCAGAAATCGCTCT 312
QY 335 ACCCGATACCGCTTATGCTCGCTTGAGAAAGATGTTATGACAGTCTGACCTGCG 394
DB 311 AGCAGCAGAGTGTCTGCTGTGAGCGGCTGCGGAGATGCTCGGCTGAGAGCCCGCA 252
QY 395 CTTCCGCGCTGCGGATTAATCTGCGGCGCGCT 427
DB 251 GCGGCGCAGGAGGTGCGGCTGAGAGCAGCT 219

RESULT 5
LOCUS DR135967/c 1217 bp mRNA linear EST 16-JUN-2005
DEFINITION 49289438 Drosophila pseudoobscura embryonic cDNA library Drosophila
ACCESSION DR135967
VERSION DR135967.1 GI:67881057
KEYWORDS EST.

SOURCE
ORGANISM Drosophila pseudoobscura
Drosophila pseudoobscura
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS Richard, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S.,
Nielsen, S., Thornton, K., Hudis, M. J., Chen, R., Weisel, R. P.,
Coutonne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Buessemaker, H. J.,
van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E.,
Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrientos, D.,
Rives, C. M., Metzker, M. L., Muzny, D. M., Scott, G., Steffen, D.,
Wheeler, D. A., Worley, K. C., Havlak, P., Durbin, K. J., Egan, A.,
Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamblon, C., Huang, Y.,
Waldron, L., Verdusco, D., Clerc-Blankenburg, K. P., Dubchak, I.,
Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W.,
Gelbart, W., Weinstein, G. M. and Gibbs, R. A.

Comparative genome sequencing of Drosophila pseudoobscura:

JOURNAL
PUMED
15632085

Chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)

COMMENT

Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephenr@bcm.tmc.edu
NCBI Trace Archive: 226715265
Insert Length: 1750 Std Error: 0.25.

FEATURES

source

1. .1217

/organism="Drosophila pseudoobscura"

/mol_type="mRNA"

/db_xref="taxon:7237"

/clone="15"

/dev_stage="0-18h embryos"

/clone_lib="Drosophila pseudoobscura embryonic cDNA library"

/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo

ORIGIN

Query Match 4.8%; Score 39.6; DB 8; Length 1217;

Best Local Similarity 57.1%; Pred. No. 4.2;

Matches 72; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 486 GATGCGGTTTCTTGGAGAGTTTGCCTGCGCCGCGGAGAGGCTGACCCGCTGA 545

DB 732 GGATGCTTTTCTTCTGTCAGGCTTCCCTCCCGCTCTTGGCCAGCCGCTGCTCC 673

QY 546 TCTGATGATGTTGAGGATTTTTCGACAGGAGGAAGATCCGTTTGGCACTCAATCC 605

DB 672 GCTCCCTTCGTTGTCGACATCTCCGATTCATGTTCTCCGCAATGCGCAATTGGTCA 613

QY 606 CGCCTT 611

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

QY 612 CAGCCT 607

DB 612 CAGCCT 607

FEATURES

source

1. .530

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:871429"

/sex="male"

/tissue_type="heart"

/dev_stage="4 weeks"

/lab_host="DH10B"

/clone_lib="Soares mouse NBMH"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5],

TGTTACCAATCTGAAGTGGAGCGCGCCGGAATTTTCTTTTCTTTTCTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT73 vector. RNA

provided by Dr. Minoru Ko, Wayne State Univ. Library

constructed and normalized by Benito Soares and M.Fatima

Bonaldo."

ORIGIN

Query Match 4.7%; Score 39.2; DB 1; Length 530;

Best Local Similarity 48.6%; Pred. No. 4.4;

Matches 107; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 230 AGAAAAATTCCTTCCGCAAGACGCTTGGCTGCAAGAACGCTTGAACCGGATACCGGCT 349

DB 125 AGAAAAATTCCTTGAAGCAAGAGGACCTTGGAGGACCAATTAATCCAGGCTAACCCG 184

QY 350 TTATGTCGCTTGGAAAGCATGTTATGACGTCGCTGACCTCGCCCTCGGCGGCGG 409

DB 185 CTCTGAGAGCCCTTGGAAAGCCCAAGACTCTCCGAATGACAACTCTCCCTTTGGAA 244

QY 410 ATTACTGGGCGGCGCTTTCGCTGTTGGAAGCGAACACTGGGGGACGCGGCGTATA 469

DB 245 ATTCACTAGATTCACCTTGGAGCTACTGAAAGCTGCTTCTGCAACATAGAGACTTA 304

QY 470 TCATTTCCCGAAAGCATGCGGTTTCTTCCGACAGTT 509

DB 305 CTTCTGGAAGAGTCGCCGGTGAATCTTCAGCTTAAGGCT 344

RESULT 7 1357 bp DNA linear GSS 13-MAY-2003
LOCUS CC306269
DEFINITION CH261-2K1.RM.1 CH261 Gallus gallus genomic clone CH261-2K1,
genomic survey sequence.
ACCESSION CC306269
VERSION CC306269.1 GI:30677710
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

AUTHORS

Kremetzki, C., Higgindocham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE

Unpublished (2003)

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

JOURNAL

Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 18200 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGAA
Class: BAC ends
High quality sequence start: 135

COMMENT

High quality sequence start: 135

FEATURES High quality sequence stop: 615.

Location/Qualifiers

1..1357
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-2X1"
/sex="female"
/cell_line="HCD001, inbred 256"
/note="Vector: pTARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match 4.7%; Score 39.2; DB 9; Length 1357;
Best Local Similarity 51.7%; Pred. No. 5.7;
Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

625 CTGCATTATGCGCAAGTTTTCAGCAAGAAAGCGGCAATGGGCGAGCTGATCGAAGCAG 684
1059 CTCTCTTACCACTCCAGCAAGAAACCAACCAATTTTGGCCCAACACACAA 1118
685 CGCCTCTGAAACCGCAAGCAAGCGGCGATTCCTCCGCGCAACATTCAGAACCGCG 744
1119 CGCGCTACCAACAAAGCAAGCAAGCGCCCTCCCGCAAAATACTACAGAACAC 1178
745 CTGATCGCGCGCTTGAACCAATTCAGAGGAAAGGAAACGCGCGCAAA 796
1179 AAACACCCACCATATCCAAACAGCGCGCAAAATATCCACCAAA 1230

RESULT 8
CC821108/c 737 bp DNA linear GSS 22-OCT-2003
DEFINITION FECP48 Uncultured human fecal virus uncultured human fecal virus
ACCESSION CC821108
VERSION CC821108.1 GI:37806908
KEYWORDS GSS.
SOURCE uncultured human fecal virus
ORGANISM Viruses; environmental samples.
REFERENCE 1 (bases 1 to 737)
Breitbart, M., Hewson, I., Felts, B., Mahaffy, J. M., Nulton, J.,
Salomon, P., and Rohwer, F.
Metagenomic analyses of an uncultured viral community from human
feces

JOURNAL J. Bacteriol. 185 (20), 6220-6223 (2003)
PUBMED 14526037

COMMENT Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr., San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: foreste@umstroke.sdsu.edu
Classes: Shocgun.

FEATURES Location/Qualifiers
1..737
/organism="uncultured human fecal virus"
/mol_type="genomic DNA"
/db_xref="taxon:239364"
/clone_lib="Uncultured human fecal virus"

ORIGIN

Query Match 4.7%; Score 39; DB 9; Length 737;
Best Local Similarity 49.3%; Pred. No. 5.5; Indels 0; Gaps 0;
Matches 102; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
262 GAGGACGAGCTTTACTCGCGCAAGGTGAGAAATTCCTTGCCGAGACGCTTGCGTG 321
|||||

Db 605 GAGTACTGCGGGAAGTGGCGGAACGAGATACAGAGACATTTGAGGCGCAAGGGCCG 546
Qy 322 CAAGAACGCTTGACCCGGATACCGCTTTATCGCTTGGAACGATGTTATGAC 381
Db 545 GAAATTAACCTGGCGCGGATCATGCTTTTCCAACTTGCGAGCGGAAGATGAG 486
Qy 382 GTTCCTGACCTCGCCCTCCGCGTGGCGGATTTACTGCGGCGCGCTTTCCGCTTGAA 441
Db 485 GACCACTACTCCCCCGCGCGGCTGCTGCGCTCATGAGAGAGCGCGCTGAGTTGAA 426
Qy 442 AGGAACTGCGGCGGAGCGCGGCTAT 468
Db 425 ACCCACTCTGGAAACCAACCTTAT 399

RESULT 9
BP210238 582 bp mRNA linear EST 15-SHP-2004
LOCUS BP210238 Sugano cDNA library, cerebellum Homo sapiens cDNA clone
DEFINITION CBL05056, mRNA sequence.
ACCESSION BP210238
VERSION BP210238.1 GI:52083129
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556

REFERENCE Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokane-dai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
TITLE Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBL05056"
/issue_type="cerebellum"
/clone_lib="Sugano cDNA library, cerebellum"

ORIGIN

Query Match 4.7%; Score 38.8; DB 3; Length 582;
Best Local Similarity 45.4%; Pred. No. 5.9;
Matches 139; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

411 TTACGCGGCGCGCTTTCGCTTGGAAGCAACATGCGGCGGCTATAT 470
Db 5 TTGATAGGGGAGAGCTCACTGATTAAGACCAACCTCGGGGGCTCAGGGCTGT 64
Qy 471 CATTTCCGAAAGCGATCGGTTTCTCGACAGTTTCCGCGCTCCCGCGGAAG 530
Db 65 CTTCCTTCCTCCCTCGCTCGCGGCTCCACACAGTTGCAACCTGCAAGAGCCCGGAGA 124
Qy 531 GCTGACCCCGCTGATCGATGATGATTCAGCGATTTTTCAGACGAGGAAGATCCGCT 590
Db 125 ACACAACCTCCCGAGAGCCAGGTCAGAGCCAAACCGTCACTGACAGTTGGCTGA 184
Qy 591 TTGCACTCAATCCCGCTTGTCGCCAGAGCTGATTAAGCAAGTTTACAGACA 650
Db 185 TGAGTCGCTGGAAGCAACCGCTGATGCTCACTGTAAGAGATTAAGATGCTGG 244
Qy 651 AAACAGCCATTTGGGCACTGATGAAACAGACCGCTCTGAAACCGCAACAGCAAG 710
Db 245 TATCAAGACTTGGTTATGTTAGTGAACAGAGAAACAACCTGAAACGCTTGAAGAGG 304
|||||

QY 711 GCGCCA 716
DB 305 GATGGA 310

RESULT 10
AM677069/c
LOCUS
DEFINITION
AM677069 469 bp mRNA linear EST 19-JUL-2000
sequence.
DGI_4_c12_b1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA

ACCESSION
AM677069
VERSION
AM677069.1 GI:7550746
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor (sorghum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 469)
Cordonnier-Pratt, M.-W., Gingle, A., Marsala, C., Sudman, M. and
Pratt, L.H.
REFERENCE
AUTHORS
TITLE
An EST database from Sorghum: dark-grown seedlings
JOURNAL
Unpublished (2000)
COMMENT
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 412
POLYA=No.

FEATURES
source
1..469
location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_1lb="Dark Grown 1 (DGI)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."

ORIGIN
Query Match 4.7%; Score 38.6; DB 1; Length 469;
Best Local Similarity 49.7%; Pred. No. 6.4;
Matches 98; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 215 TGTGAGAGGAGCATTTGAGCAAGGTCTGCCGTATATACCGTATTGAGACGATT 274
DB 205 TCTGGAAGTGTGTTGGGAGACGAGAACTGATGTTCCATGTTTTCGAGACCTAACT 146
QY 275 TACTGCGGAGAGTGAAGAAAATTCCTTGCAGAGACGCTTGCTGCAAGACGCTTG 334
DB 145 TGCTCCGAGACGGGTCTTAACAAGCTTGGCCGATGTCGTTAGGCCCAAGATCCCTCT 86
QY 335 ACCCGGATACCGGCTTTATCGTCCGCTTGGAAGAGATGTTATGACCGCTCGACCTGCG 394
DB 85 AGCAGCAGAGTGTTCCTGCTTGCGAGCGGGGTGGCGGAAGTCTCGCGCTGAGGCCCGA 26
QY 395 CCTCGGCGGTGCGGAT 411
DB 25 GCGGCCGACAGGAGGT 9

RESULT 11
BZ577887/c
LOCUS
BZ577887 775 bp DNA linear GSS 17-DEC-2002

DEFINITION
mesh2_5621.x1 mesh Pseudomonas aeruginosa genomic clone mesh2_5621,
genomic survey sequence.
ACCESSION
BZ577887
VERSION
BZ577887.1 GI:27212948
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 775)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 357145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1..775
location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="mesh2_5621"
/clone_1lb="mesh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 4.7%; Score 38.6; DB 9; Length 775;
Best Local Similarity 49.3%; Pred. No. 7.3;
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 344 CCGCCTTATGTCGCGCTTGGAAGAGATGTTATGCAAGTCCGACCTCGCCCTCCGCG 403
DB 546 CCGGCGTATGCGCTTGAGCTGCTGCTGCTGCGCGCGCTGCGACAGTCAACGG 487
QY 404 TGGCGATTTACTGCGGCGCGCTTCCGCTGTTGGAAGCGAACACTGCGGCGCGCG 463
DB 486 TGCTGAGACTACCTGAGGAGATGCGCGGCGCTGAGATGGGAAACCGAAGACCTGCG 427
QY 464 GCTATATCATTTCCCGAAGACGATGCGGCTTTTCTGAGACAGGTTGCGCCTGCGCG 523
DB 426 AGCGCGCTTACCCCGCAGGCGCATGCGGTTTCGCTCGGACCCGCGTGTGCGCGCGC 367
QY 524 CCGAAGGCGTCAACCCGCTGATCT 548
DB 366 GCATGCGGAGCAGGCGCTCGAAT 342

RESULT 12
CNSOP4HI
LOCUS
DEFINITION
CNSOP4HI 920 bp mRNA linear HTC 05-JUL-2005
Tetraodon nigroviridis full-length cDNA.
CNSOP4HI
ACCESSION
CR655152
VERSION
CR655152.2 GI:56246766
KEYWORDS
HTC; cDNA; full-length; full-length cDNA; Tetraodon nigroviridis.
Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE
AUTHORS
Jaillon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thumann, N.,
Maucci, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A.,
Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C.,

TITLE
 JOURNAL
 PUBLISHED
 REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 COMMENT
 FEATURES
 SOURCE
 ORIGIN
 Query Match
 Best Local Similarity 51.1%; Score 38.4; DB 4; Length 920;
 Matches 90; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 Oy 631 TATGCCAAGTTTACGACCAAAACGCGCATTTGGCAGCTGATGAAACGACCGCTTC 690
 Db 236 TCTGCTGGAGAAACGAGAAAGGCGCGAAGGGCGACGGGACAGCTGATATCATTC 295
 Oy 691 CTGAACCGCAACGCAAGAGGCGGATTCCTCCCGCCACACATTCACAAACCGCGTGTTC 750
 Db 296 CTCACACCAACCTGCAACCTCTCGCTTTCTCAAAAAGTCTTATATGTTACGTCAA 355
 Oy 751 CGCGCTTGACCAAAATCAGCAGGGAAGGAAAACCGCGGCAAAAGCGCGCAACA 806
 Db 356 GACTACATGAAAGCATCATAGAGCCACTGTGAGGAACACAAACCCAGCGCGCATCA 411
 RESULT 13
 CN206175 569 bp mRNA linear EST 30-APR-2004
 LOCUS CN206175
 DEFINITION Tor6607 Gametophyte rehydration library Tortulla ruralis cDNA, mRNA
 sequence.
 ACCESSION CN206175
 VERSION CN206175.1 GI:46902906
 KEYWORDS EST.
 SOURCE Tortulla ruralis
 ORGANISM Tortulla ruralis
 Buariyoda; Vitridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Dictyonidae; Potillales; Potillaceae; Tortulla.
 1 (bases 1 to 569)
 Oliver,M.J., Dowd,S.E., Zaragoza,J., Mauge,S.A. and Payton,P.R.
 The rehydratation transcriptome of the desiccation-tolerant bryophyte
 Tortulla ruralis: transcript classification and analysis
 BMC Genomics 5 (1), 89 (2004)
 15546486
 Contact: Oliver Melvin J
 Plant Stress Lab
 USDA-ARS
 3810 4th St, Lubbock, TX 79415, USA
 Tel: 806-749-5560
 Fax: 806-723-5272

Email: moliver@lbk.ars.usda.gov
 PCR Primers
 FORWARD: GTTTTCCAGTACAC
 BACKWARD: CAGAAACAGCTATGAC.
 location/Qualifiers
 1. 569
 /organism="Tortula ruralis"
 /mol_type="mRNA"
 /db_xref="taxon:38588"
 /clone_lib="Gametophyte rehydration Library"
 /note="Organ: Green Gametophyte; Vector: pSport1; Site_1:
 Salt; Site_2: NoLi"

ORIGIN

Query Match 4.6%; Score 38.2; DB 7; Length 569;
 Best Local Similarity 48.4%; Pred.No. 8.9;
 Matches 137; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 319 CTGCAAGAACCTTTGACCCGGAATACGCTTTATCGTCGCTTGAAACGATGTTATG 378
 |||||
 Db 204 CTGGAGATCACTTTTGGACAAACCAACAGTATTTCTTCCGCCAAGAAATGATGTTGTG 263
 |||||
 QY 379 CACGCTCGACCTGCGCCCTCCGGGGTGGCGGATTAATGCGGGGCGCGCTTTCCGCTGTG 438
 |||||
 Db 264 GAGTTCATGACCGTGGGAGACGGGACGAGCAATTCACAGGCTCTCCACACAGAGAG 323
 |||||
 QY 439 GA---AAGCGAACACTGGGGGACCGCGGGCTATATCAATTTCCGAAAAAGCATGCGGTTT 495
 |||||
 Db 324 GACTTCGGCGGGGCGATATGTGGCGCGCGCGCTCAGCTCGGCCGACATTCACACCGCGGCC 383
 |||||
 QY 496 TTCCTGACAGGTTTGGCGCCCTGCGCGCGGAGAGGGCTGCACCCCGTCGATCTGATGATG 555
 |||||
 Db 384 ATGCAGAGGCGCGGTGACGCTCACCCCAAGTACAGCATGAAAGCAGGCGCATGCAAGAGATG 443
 |||||
 QY 556 TTCAGCGATTTTTCGACAGGAGAGAAATGCCGCTTTGCCAGC 598
 |||||
 Db 444 TTCAGGGCTCGACAGCAAGGGGCTACATGATGAGACACCATC 486
 |||||

RESULT 14

LOCUS CF687694 553 bp mRNA linear EST 16-AUG-2004
 DEFINITION CCACV03TR C.neoformans strain JEC21 Cryptococcus neoformans var.
 neoformans cDNA clone CCACV03, mRNA sequence.
 ACCESSION CF687694
 VERSION CF687694.1 GI:41541853
 KEYWORDS EST.
 SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
 var. neoformans)
 ORGANISM Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.
 1 (bases 1 to 553)
 REFERENCE 1
 AUTHORS Loftus,B.
 TITLE End sequencing of clones from a Full length enriched, normalized
 JEC21 cDNA library
 JOURNAL Unpublished (2003)
 COMMENT Other_ESTs: CCACV03TF
 Contact: Brendan Loftus
 The Institute for Genomic Research (TIGR; www.tigr.org)
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: b1loftus@tigr.org
 Seg primer: TR.
 FEATURES
 source Location/Qualifiers
 1. 553
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCACV03"

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:17:47 ; Search time 296.232 Seconds
(without alignments)
4968.475 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacaccagcttaccag.....tcattgtgccttccaataa 828

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/R COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	726	87.7	5859	2	US-08-312-387B-1
2	726	87.7	5859	2	US-08-312-387B-1
3	726	87.7	5859	2	US-08-683-426-1
4	726	87.7	5859	2	US-08-683-426-1
5	726	87.7	5859	2	US-08-683-426-1
6	726	87.7	5859	2	US-08-683-458-1
7	726	87.7	5859	2	US-08-683-458-1
8	726	87.7	5859	2	US-08-878-360-1
9	726	87.7	5859	2	US-08-878-360-1
10	726	87.7	5859	3	US-08-878-140B-1
11	726	87.7	5859	3	US-09-333-412-1
12	726	87.7	5859	3	US-09-333-412-1
13	726	87.7	5859	3	US-09-338-943-1
14	726	87.7	5859	3	US-10-007-267A-1
15	726	87.7	5859	3	US-10-007-267A-1
16	86.6	10.5	1830121	3	US-09-557-884-1
17	86.6	10.5	1830121	3	US-09-643-990A-1
18	73.2	8.8	768	3	US-10-158-865-1
19	73.2	8.8	65792	3	US-09-540-236-487
20	57	6.9	1830121	3	US-09-557-884-1
21	57	6.9	1830121	3	US-09-643-990A-1
22	57	6.9	1830121	3	US-10-158-865-1
23	39.4	4.8	4455	3	US-09-902-540-7119
24	39.4	4.8	4458	3	US-09-902-540-650

25	39	4.7	399	3	US-09-902-540-2937	Sequence 2937, App
26	39	4.7	23417	3	US-09-902-540-1207	Sequence 1207, App
27	38.6	4.7	753	3	US-09-252-991A-10290	Sequence 10290, A
28	38.6	4.7	1260	3	US-09-252-991A-10493	Sequence 10493, A
29	38.6	4.7	1635	3	US-09-252-991A-10567	Sequence 10567, A
30	38.6	4.7	1743	3	US-09-252-991A-10674	Sequence 10674, A
31	37.2	4.5	1401	4	US-09-605-703B-543	Sequence 543, App
32	36.8	4.4	31422	3	US-09-914-286-2	Sequence 2, App1
33	36	4.3	1721	2	US-07-688-352C-13	Sequence 13, App1
34	36	4.3	1721	2	US-08-474-375C-13	Sequence 13, App1
35	36	4.3	1721	3	US-09-146-249A-13	Sequence 13, App1
36	36	4.3	1721	3	US-08-206-188B-13	Sequence 13, App1
37	36	4.3	1721	6	PCT-US91-02714-13	Sequence 13, App1
38	35.8	4.3	1259	2	US-08-997-080-123	Sequence 123, App
39	35.8	4.3	1259	2	US-08-997-362-123	Sequence 123, App
40	35.8	4.3	1259	3	US-09-095-855-123	Sequence 123, App
41	35.8	4.3	1259	3	US-09-324-542-123	Sequence 123, App
42	35.8	4.3	1259	3	US-09-205-426-123	Sequence 123, App
43	35.8	4.3	2013	2	US-08-997-080-176	Sequence 176, App
44	35.8	4.3	2013	2	US-08-997-362-176	Sequence 176, App
45	35.8	4.3	2013	3	US-09-095-855-176	Sequence 176, App

ALIGNMENTS

RESULT 1
US-08-312-387B-1
; Sequence 1, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5859 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Neisseria gonorrhoeae
; STRAIN: F62
; FEATURE:
; NAME/KEY: CDS

LOCATION: 1..381
FEATURE: CDS
NAME/KEY: 445..1491
LOCATION: 445..1491
FEATURE: CDS
NAME/KEY: 2342..3262
LOCATION: 2342..3262
FEATURE: CDS
NAME/KEY: 3322..4335
LOCATION: 3322..4335
FEATURE: CDS
NAME/KEY: 4354..5196
LOCATION: 4354..5196
US-08-312-387B-1

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGAGGAGGAGCAATTGCCGAT 60
1491 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGAGGAGGAGCAATTGCCGAT 1550
61 ACCCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
1551 ACCCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1610
121 CTGGAACAGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
1611 CTGGAACAGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1670
181 GTGGAAGAAAGGCTCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
1671 GTGGAAGAAAGGCTCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1730
241 CTGCGCTATATACAGCTTAGCTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
1731 GTACCGTATATACAGCTTAGCTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1790
301 CTGCGCAAGAGGCTCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
1791 CTGCGCAAGAGGCTCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1850
361 TTGGAAGAGATGTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
1851 TTGGAAGAGATGTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1910
421 CGGCGCTTCCGCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
1911 CGGCGCTTCCGCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1970
481 AAAGCGATGCGGTTTCTGAGCAGGTTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 540
1971 AAAGCGATGCGGTTTCTGAGCAGGTTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 2030
541 GTGCAATGATATAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
2031 GTGCAATGATATAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2090
601 AATCCGCGCTTGTGCGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
2091 AATCCGCGCTTGTGCGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2150
661 TTGGGAGGCTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
2151 TTGGGAGGCTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2210
721 CCGGCAACAGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
2211 CCGGCAACAGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
781 GAAAAAGCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 814

Db 2271 GAAAAAGCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2304

RESULT 2
US-08-312-387B-7
Sequence 7, Application US/08312387B
Patent No. 5545553

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,387B
FILING DATE: July 7, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 5859 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Neisseria gonorrhoeae

STRAIN: F62

FEATURE:

NAME/KEY: CDS

LOCATION: 1491..2330

US-08-312-387B-7

Query Match 87.7%; Score 726; DB 2; Length 5859;

Best Local Similarity 93.2%; Pred. No. 2,8e-216;

Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGAGGAGGAGGAGCAATTGCCGAT 60
1491 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1550
61 ACCCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
1551 ACCCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1610
121 CTGGAACAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
1611 CTGGAACAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1670
181 GTGGAAGAAAGGCTCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
1671 GTGGAAGAAAGGCTCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1730

OY	241	CTGCAGTATACACCGTATTGGAGACAGCGTTTACTCTGGCGAAGGTGAAGAAAATTC	300
Db	1731	GTACCGTATATCCGCGTATTGGAATGATGTCTTACTGGCGGAAGGCGAGCAATTTC	1790
OY	301	CTTGCAGAAACGCTTGGCTCGAAMACGCTTTTGACCGGATATACCGCTTATGTCGCG	360
Db	1791	CTTGCAGAAAGTACTTGGCTCGAAMAAACGCTTTGACCCCGATTCCGCTTGTGTCGCG	1850
OY	361	TTGGAAACGATGTTTATGACGTCCTGACCTTCGCCCTTCGGCGTGGCGGAATTACTCGCGG	420
Db	1851	TTGGAAACGATGTTTATGACGTCCTGACCTTCGCCCTTCGGCGTGGCGGAATTACTCGCGG	1910
OY	421	CGCGCTTTCGCTGTGGAAAGGAAACAATCGGGGGAAGGGGGGTATATCATTTCCCGA	480
Db	1911	CGCGCTTTCGCTGTGGAAAGGAAACAATCGGGGGAAGGGGGGTATATATTTTCCCGA	1970
OY	481	AAAGCGATGCGGTTTCTTGGAACAGTTTCCGCGCTCGCGCGCCGGAAGGGCTCAACCC	540
Db	1971	AAAGCGATGCGGTTTCTTGGAACAGTTTCCGCGCTTTCGCGCGCGGAAGGGCTCAACCT	2030
OY	541	GTGCATCTGATGATGTTTCAAGCGATTTTTTCAGACAGGAAGGAATGCGGTTTGCAGCTC	600
Db	2031	GTGCATTTGATGATGTTTTCGCAACCCCTGACGACAGGAAGGAATGCGGTTTGCAGCTC	2090
OY	601	AATCCCGGCTTGTGGGCCCAAGAGCTGTATATGCGAGTTTTCAGACCAAAACAGCGCA	660
Db	2091	AATCCCGGCTTGTGGGCCCAAGAGCTGTATATGCGAGTTTTCAGACCAAAACAGCGCA	2150
OY	661	TTGGGCAACCTGATGGAACAGACCGGCTCTCTGAAACGCAAAACAGCAAAAGGCGGATTC	720
Db	2151	TTGGGCAACCTGATGGAACAGACCGGCTCTGAAACGCAAAACAGCAATGCGGGAATTC	2210
OY	721	CCCGGCAACATTTCAAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG	780
Db	2211	CCCGGCAACATTTCAAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG	2270
OY	781	GA AAAACGCGCGCAAAAGCGCGAACAAGTTCAATG 814	
Db	2271	GA AAAACGCGCGCAAAAGCGCGAACAAGTTCAATCG 2304	

RESULT 3
 US-08-663-426-1
 Sequence 1, Application US/08663426
 Patent No. 5705357
 GENERAL INFORMATION:
 APPLICANT: Gotschlich, Emil C.
 TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
 TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/663,426
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/312,387
 FILING DATE: September 26, 1994
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742

```

? REFERENCE/DOCKET NUMBER: 600-1-095B
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201 487-5800
? TELEFAX: 201 343-1684
? TELEX: 133521
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5859 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Neisseria gonorrhoeae
? STRAIN: F62
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..381
? OTHER INFORMATION: glyc (glycyl tRNA synthetase beta chain)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 445..1491
? OTHER INFORMATION: 1gta
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 2342..3262
? OTHER INFORMATION: 1gtc
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 3322..4335
? OTHER INFORMATION: 1gtd
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 4354..5196
? OTHER INFORMATION: 1gte
?
? US-08-663-426-1

```

Query Match	87.7%	Score 726	DB 2	Length 5859			
Best Local Similarity	93.2%	Pred. 2.8e-216					
Matches 759	Conservative 0	Mismatches 55	Indels 0	Gaps 0			
QY	1	ATGCAAAACCA	CGTTATCA	GCTTAGCTTCGCGCGCAGAA	CGCAGGCGCA	CATTGCCGAT	60
Db	1491	ATGCAAAACCA	CGTTATCA	GCTTAGCTTCGCGCGCAGAA	CGCAGGCGCA	CATTGCCGCA	1550
QY	61	ACCTTCGGCA	GGCAGCGGCA	ATCCGTTTCA	GTTTTCGAC	GCACCTGATGCCGTC	120
Db	1551	ACCTTCGGCA	GTCCCGGCAT	CCCGTTCCAGTTTTC	GACGCAC	GTGATGCCGTC	1610
QY	121	CTGGAA	CAGGCAAT	AGGCGGGAAC	TCGTCTCCCGGCTGTG	CGGCGCACCCCTA	180
Db	1611	CTGGAA	CAGGCAAT	AGGCGGGAAC	TCGTCTCCCGGCTGTG	CGGCGCACCCCTA	1670
QY	181	GTGAAAAA	AGCGCTGCTTTAT	TGACGCAC	GCCTGATTTG	TGGAAGCAG	240
Db	1671	GTGAAAAA	AGCGCTGCTTTAT	TGACGCAC	GCCTGATTTG	TGGAAGCAG	1730
QY	241	CTGCCGTA	TATCAC	CGTATTTG	AGGACG	AGTTTTACTCGGCGA	300
Db	1731	GTACCCGTA	TATCGCGCTAT	TTTGAAGATG	CTTACTCGGCGA	AGCGCGGAC	1790
QY	301	CTTGC	CGAAGCGCTT	GGCTGCAAG	ACGCTTTGA	ACC	360
Db	1791	CTTGC	CGAAGTA	CTTGGCTG	CAAGAA	CGCTTTGA	1850
QY	361	TTGAAA	ACGATGTTTAT	TGACAGCT	CTACCTCGC	CTCCGCGGTG	420
Db	1851	TTGAAA	ACGATGTTTAT	TGACAGCT	CTACCTCGC	CTCCGCGGTG	1910
QY	421	CGCGC	CTTTCGCTG	TTGAAA	ACGAA	CAC	480

Db 1911 CGCGCTTTCCGCTTTTGGAAAGCGAACCTGCGGAGCGCGGCTATATTATTTCCCGA 1970
Qy 481 AAGCGATGCGGTTTTCCTGGAAGATTTGCGCCCTGCGCCGCAAGGAGCTGACCCC 540
Db 1971 AAGCGATGCGGTTTTCCTGGAAGATTTGCGCCCTGCGCCGCAAGGAGCTGACCCC 2030
Qy 541 GTGATCTGATGATGTTTCAGCGGATTTTTCGACAGGAAAGAAATGCGGTTTCCGAGTC 600
Db 2031 GTGATCTGATGATGTTTCAGCGGATTTTTCGACAGGAAAGAAATGCGGTTTCCGAGTC 2090
Qy 601 AATCCGCTTTGTCGCGCCCAAGAGCTGATTCATTCAGAGTTTCACGACCAAAACAGCGCA 660
Db 2091 AATCCGCTTTGTCGCGCCCAAGAGCTGATTCATTCAGAGTTTCACGACCAAAACAGCGCA 2150
Qy 661 TTGGGAGCGCTGATTCGACAGCAGCGCTTCGACCGCAAGCAAGAGGCGGCTTC 720
Db 2151 TTGGGAGCGCTGATTCGACAGCAGCGCTTCGACCGCAAGCAAGAGGCGGCTTC 2210
Qy 721 CCGGCAACACATTTCAAACACCGCTGATTCGCGGCTTCGACCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTTCAAACACCGCTGATTCGCGGCTTCGACCAAAATCAGGAGAAAG 2270
Qy 781 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 814
Db 2271 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 2304

RESULT 4

US-08-683-426-7
; Sequence 7, Application US/08683426
; Patent No. 5705367

; GENERAL INFORMATION:
; APPLICANT: GoeSchlich, Emil C.

; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:
; ADDRESS: Klauder & Jackson

; STREET: 411 Hackensack Avenue
; CITY: Hackensack

; STATE: New Jersey
; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,426

; FILING DATE:
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387

; FILING DATE: September 26, 1994
; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-0958

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800

; TELEFAX: 201 343-1684
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 5859 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: both
; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

; ANTI-SENSE: NO
;

;/ ORIGINAL SOURCE:
;/ ORGANISM: Neisseria gonorrhoeae
;/ STRAIN: F62
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: 1491..2330
;/ OTHER INFORMATION: 19cB
;/ US-08-683-426-7

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1 AAGCAAAACCAAGTATGAGCTTACCTCCGCGGAGAAACGAGGCGGCAATTCGCGAT 60
Db 1491 ATGCAAAACCAAGTATGAGCTTACCTCCGCGGAGAAACGAGGCGGCAATTCGCGCA 1550
Qy 61 ACCCTCGGAGGCAAGGATCCGTTTCAAGTTTTCAGACGATGATCCGCTGAAAG 120
Db 1551 ACCCTCGGAGGCAAGGATCCGTTTCAAGTTTTCAGACGATGATCCGCTGAAAG 1610
Qy 121 CTGGAACGGAATGCGGGAAGCTGTCGCGGCTTGTGCGGCAACCCCTATTGAGCGGA 180
Db 1611 CTGGAACGGAATGCGGGAAGCTGTCGCGGCTTGTGCGGCAACCCCTATTGAGCGGA 1670
Qy 181 GTGAAAAAGCTGCTTTATGAGCCAGCGCTGATTTGTGAAAGCAGGCAATTGACGAGGT 240
Db 1671 GTGAAAAAGCTGCTTTATGAGCCAGCGCTGATTTGTGAAAGCAGGCAATTGACGAGGC 1730
Qy 241 CTGCGGTATATCACCGTATTTGAGGACGACGTTTACTCGGGAAGGAGGAAAAATTC 300
Db 1731 GTACGTTATTCGCGGTATTTGAGGATGATCTTACTCGGGAAGGCGGAGATTC 1790
Qy 301 CTTCGCAAGAGCTTTCGCTGCAAGAGCTTTGACCCGGAATCCGCTTTATGTCGCC 360
Db 1791 CTTCGCAAGAGCTTTCGCTGCAAGAGCTTTGACCCGGAATCCGCTTTATGTCGCC 1850
Qy 361 TTGAAAGAGATGTTATGACAGTCTGACCTCGCCTCGGCGGAGGAGGATTAATGCGAGG 420
Db 1851 TTGAAAGAGATGTTATGACAGTCTGACCTCGCCTCGGCGGAGGAGGATTAATGCGAGG 1910
Qy 421 CCGGCTTTCCGCTGTTGAAAGGCAACATGCGGGAAGCGGCGTATATCATTTCCCGA 480
Db 1911 CCGGCTTTCCGCTGTTGAAAGGCAACATGCGGGAAGCGGCGTATATTAATTTCCCGA 1970
Qy 481 AAGGATGCGGATTTTCTGGAAGGATTTGCGCCCTGCGCGGAGAGGCTGCAACCC 540
Db 1971 AAGGATGCGGATTTTCTGGAAGGATTTGCGCCCTGCGCGGAGAGGCTGCAACCC 2030
Qy 541 GTGATCTGATGATGTTTCAGAGATTTTTCGACAGGAAAGAAATGCGGTTTCCAGCTC 600
Db 2031 GTGATCTGATGATGTTTTCAGAGATTTTTCGACAGGAAAGAAATGCGGTTTCCAGCTC 2090
Qy 601 AATCCGCTTTGTCGCGCCCAAGAGCTGATTCAGAGTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCGCTTTGTCGCGCCCAAGAGCTGATTCAGAGTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGAGCGCTGATTCGACAGCAGCGCTTCGAAACGCAAAAGCGGAGATTC 720
Db 2151 TTGGGAGCGCTGATTCGACAGCAGCGCTTCGAAACGCAAAAGCGGAGATTC 2210
Qy 721 CCGGCAACACATTTCAAACACCGCTGATTCGCGGCTTCGACCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTTCAAACACCGCTGATTCGCGGCTTCGACCAAAATCAGCAGGAAAG 2270
Qy 781 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 814
Db 2271 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 2304

RESULT 5

US-08-683-458-1
; Sequence 1, Application US/08683458

Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULAR TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1..381
OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
FEATURE:
NAME/KEY: CDS
LOCATION: 445..1491
OTHER INFORMATION: 1gfa
FEATURE:
NAME/KEY: CDS
LOCATION: 2342..3262
OTHER INFORMATION: 1gfc
FEATURE:
NAME/KEY: CDS
LOCATION: 3322..4335
OTHER INFORMATION: 1gfd
FEATURE:
NAME/KEY: CDS
LOCATION: 4354..5196
OTHER INFORMATION: 1gre
US-08-683-458-1
Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGCAAAACCACTTATCAGCTTACGCTCCGCGGAGAACGAGGCGACATTGCCGAT 60

Db 1491 ATGCAAAACCACTTATCAGCTTACGCTCCGCGGAGAACGAGGCGACATTGCCGCA 1550
QY 61 ACCTTGCGAGGACGACATCCCGTTTCAGTTTTCAGACGACTGATGCCGTGTAAGG 120
Db 1551 ACCTTGCGAGGACGACATCCCGTTTCAGTTTTCAGACGACTGATGCCGTGTAAGG 1610
QY 121 CTGGAACAGGCAATGCGGAACTCCGCTTCGCTTCGCGGACCCCTATTGACCGGA 180
Db 1611 CTGGAACAGGCAATGCGGAACTCCGCTTCGCTTCGCGGACCCCTATTGACCGGA 1670
QY 181 GTGAAAAAGCCCTGTTTATGAGCCAGCGCGTATTGTGGAAGGACATTGACGAAGT 240
Db 1671 GTGAAAAAGCCCTGTTTATGAGCCAGCGCGTATTGTGGAAGGACATTGACGAAGG 1730
QY 241 CTGCGGTATATACCGGTATTTAGAGCAGCTTTTATCTCGCGAAGTGAAGAAAAATTC 300
Db 1731 GTACCGTATATGCGCGTATTTGAAATGATGCTTACTCGGCGAAGGCGGAGCAGTTTC 1790
QY 301 CTTCGCGAAGACGCTTGCTGCAAGAACGCTTTGACCCGGAATACGCCCTTATGTCGCC 360
Db 1791 CTTCGCGAAGATACCTTGCTGCAAGAACGCTTTGACCCGGAATACGCCCTTATGTCGCC 1850
QY 361 TTGAAAACGATGTTTATGACGCTGACGCTGCGGCTCGGCGGTGCGGATTAATGCGGG 420
Db 1851 TTGAAAACGATGTTTATGACGCTGACGCTGCGGCTCGGCGGTGCGGATTAATGCGGG 1910
QY 421 CGCGCCTTTCGCTGTGGAAGGCAACCTGCGGAGCGCGGCTATATCAATTTCCCGA 480
Db 1911 CGCGCCTTTCGCTGTGGAAGGCAACCTGCGGAGCGCGGCTATATCAATTTCCCGA 1970
QY 481 AAGGCAATGCGGTTTCTGCAAGGTTTTCGCGGCTGCGGCGGCAAGGCTGCAAGG 540
Db 1971 AAGGCAATGCGGTTTCTGCAAGGTTTTCGCGGCTGCGGCGGCAAGGCTGCAAGG 2030
QY 541 GTGCAATGATGATGTTGACGATTTTTCGACAGGGAAGAAATGCGGTTTTCGCAAGG 600
Db 2031 GTGCAATGATGATGTTGACGATTTTTCGACAGGGAAGAAATGCGGTTTTCGCAAGG 2090
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGCAACCAACAGCGCA 660
Db 2091 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGCAACCAACAGCGCA 2150
QY 661 TTGGGCAAGCTTATGGAACAGACCGCTTCTGAAACCGCAACAGCAAGGCGCATTC 720
Db 2151 TTGGGCAAGCTTATGGAACAGACCGCTTCTGAAACCGCAACAGCAAGGCGCATTC 2210
QY 721 CCCGCAACACATTCAAACACCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG 780
Db 2211 CCCGCAACACATTCAAACACCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG 2270
QY 781 GAAAAAGCCGCGCAAAAGGCGGCAACGTTCAATG 814
Db 2271 GAAAAAGCCGCGCAAAAGGCGGCAACGTTCAATG 2304
RESULT 6
US-08-683-458-7
Sequence 7, Application US/08683458
Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1491..2330
OTHER INFORMATION: 15TB
US-08-683-458-7

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACGAGTATGAGTTAGCTTCCGCGGAGAAACGAGGCGGACATTGCCAT 60
1491 ATGCAAAACGAGTATGAGTTAGCTTCCGCGGAGAAACGAGGCGGACATTGCCGA 1550
61 ACCCTCGGAGGACGAGCATCCCGTTTCAGTTTTCGACGACATGCGCTGAAAG 120
1551 ACCCTCGGAGTCCGGGATCCCGTTTCAGTTTTCGACGACATGCGCTGAAAG 1610
121 CTGGAACAGGCAATGGCGGAATCGTCCCGGCTGTGCGGCAACCCCTATTGAGCGGA 180
1611 CTGGAACAGGCAATGGCGGAATCGTCCCGGCTGTGCGGCAACCCCTATTGAGCGGA 1670
181 GTGAAAAAGCTGCTTATGAGCAAGCGCTATTGTGGAAGCAGGCAATGGAAGGT 240
1671 GTGAAAAAGCTGCTTATGAGCAAGCGCTATTGTGGAAGCAGGCAATGGAAGGT 1730
241 CTGCGGTATATCAACGATTTTGAAGACGATTTTACTCGGGAAGGTGAGAAAAATTC 300
1731 GTACCGTATATCGCGTATTTGAAGATGATCTTAATCGGGAAGGTGAGAAAAATTC 1790
301 CTGCGGAAGACGCTTGGCTGCAAGACGCTTGAACCGGGAATACCGCTTTATCGTCGC 360
1791 CTGCGGAAGACGCTTGGCTGCAAGACGCTTGAACCGGGAATACCGCTTTATCGTCGC 1850
361 TTGAAACGATTTTATGACGCTGCACTTCGCTCGGCGGTGAGCGGATTAATGCGG 420
1851 TTGAAACGATTTTATGACGCTGCACTTCGCTCGGCGGTGAGCGGATTAATGCGG 1910
421 CGCGCTTTCGCTGTGGAAGCGAAGCCTGGGGAAGCGGCGCTATATCAATTTCCGA 480
1911 CGCGCTTTCGCTGTGGAAGCGAAGCCTGGGGAAGCGGCGCTATATCAATTTCCGA 1970

481 AAAGCATCGGTTTCTGAGACAGGTTTCCGCGCTGCGCGGAGGCGTGCACCC 540
1971 AAGGCGATGCTTTTCTTGGACAGGTTTCCGCTTTCGCGCGGAGGCGTGCACCC 2030
541 GTGCATCTGATGATGTTGAGGATTTTTCAGACGGAAGAAAGCGGTTTGCAGCTC 600
2031 GTGCATCTGATGATGTTGAGGATTTTTCAGACGGAAGAAAGCGGTTTGCAGCTC 2090
601 AATCCGCTTGTGCGCCCAAGAGCTGATTTATGCAAGTTTCAAGACCAAGACGCA 660
2091 AATCCGCTTGTGCGCCCAAGAGCTGATTTATGCAAGTTTCAAGACCAAGACGCA 2150
661 TTGGCAGCTGATGAAACAGACCGCTCTGAAACGCAACGCAAGCGGATTC 720
2151 TTGGCAGCTGATGAAACAGACCGCTCTGAAACGCAACGCAATGCGGATTC 2210
721 CCGGCAACATTCAAACACCGCTGATCCGCGCTTGAACCAATCAGCAGGAAAG 780
2211 CCGGCAACATTCAAACACCGCTGATCCGCGCTTGAACCAATCAGCAGGAAAG 2270
781 GAAAAACGCGCAAGGCGGAAACAGTTCAATTG 814
2271 GAAAAACGCGCAAGGCGGAAACAGTTATCG 2304

RESULT 7
US-08-678-360-1
Sequence 1, Application US/08878360
Patent No. 5945322
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,360
FILING DATE: 18-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/683,426
FILING DATE:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1..381
OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
FEATURE:
NAME/KEY: CDS
LOCATION: 445..1491
OTHER INFORMATION: 1gca
FEATURE:
NAME/KEY: CDS
LOCATION: 2342..3262
OTHER INFORMATION: 1gfc
FEATURE:
NAME/KEY: CDS
LOCATION: 3322..4335
OTHER INFORMATION: 1gfd
FEATURE:
NAME/KEY: CDS
LOCATION: 4354..5196
OTHER INFORMATION: 1gcb
US-08-878-360-1

Query Match 87.7%; Score 726; DB 2; Length 5859;

Best Local Similarity 93.2%; Pred. No. 2.8e-216; Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTTATGAGCTTACGCTTCCGCGCAAGAACGCGGCGCATTTGCCGAT 60
1491 ATGCAAAACCAAGTTATGAGCTTACGCTTCCGCGCAAGAACGCGGCGCATTTGCCGAT 1550
61 ACCTTGGCGAGGACGCGCATCCCGTTTCACTTTTGGAGCATGATGCCGTCTGAAAG 120
1551 ACCTTGGCGAGGACGCGCATCCCGTTTCACTTTTGGAGCATGATGCCGTCTGAAAG 1610
121 CTGGAACAGGCAATGGCGGAACTGTCCTCCGCGCTGTCGCGCGGACCCCTATTGAGCGGA 180
1611 CTGGAACAGGCAATGGCGGAACTGTCCTCCGCGCTGTCGCGCGGACCCCTATTGAGCGGA 1670
181 GTGGAAGAAAGCTGCTTTATGAGCAAGCGCGATTTGGAAGAGGCAATTGAGCAAGGT 240
1671 GTGGAAGAAAGCTGCTTTATGAGCAAGCGCGATTTGGAAGAGGCAATTGAGCAAGGT 1730
241 CTGCGGTATATCAAGCTTATTTGAGAGCAAGTTTATCTCGGCGAAGGTGAGAAAAATTC 300
1731 GTACCGTATATCGCCGTATTTGAGAGATGATGCTTACTCGGCGAAGCGGAGGAGTTC 1790
301 CTGCGGAGAGCGCTTGGCTGCAAGAACGCTTTGACCGGATACCGCTTTATGTCGCG 360
1791 CTGCGGAGAGATCTTGGCTGCAAGAACGCTTTGACCGGATACCGCTTTATGTCGCG 1850
361 TTGGAAGAGATTTATGAGCAAGCTGCAAGCTGCGCGCTGCGGCGATTTACGCGG 420
1851 TTGGAAGAGATTTATGAGCAAGCTGCAAGCTGCGCGCTGCGGCGATTTACGCGG 1910
421 CGCGCTTTCGCTGTTGGAAGAGCAACTGCGGAGGACGCGGCTATATCAATTTCCGA 480
1911 CGCGCTTTCGCTTTCGTTGGAAGAGCAACTGCGGAGGACGCGGCTATATTTCCGA 1970
481 AAAGGAGATGCGGTTTTCTGGAAGAGTTTGGCGCCCTGCGCGGAGGCTGACCCC 540
1971 AAAGGAGATGCGGTTTTCTGGAAGAGTTTGGCGCCCTGCGCGGAGGCTGACCCC 2030
541 GTGATCTGATGATTTGAGCAAGCTTTTTCAGAGGAGAGATGCGGCTTTCGAGCTC 600
2031 GTGATCTGATGATTTGAGCAAGCTTTCAGAGGAGAGATGCGGCTTTCGAGCTC 2090
601 AATCCGCTTTCGCGCGCAAGAGCTGCAATATGCAAGTTTTCAGCAACCAAAACAGCGCA 660
2091 AATCCGCTTTCGCGCGCAAGAGCTGCAATATGCAAGTTTTCAGCAACCAAAACAGCGCA 2150

661 TTGCGAGCGCTGATGCAACAGCGCGCTCTGTAACCGCAAGCAAGGCGCGATTC 720
2151 TTGCGAGCGCTGATGCAACAGCGCGCTCTGTAACCGCAAGCAAGGCGCGATTC 2210
721 CCGCGCAACATTCATAAGCAGCGCTGATCCGCGCTTGCACAAATACAGCAGGAAG 780
2211 CCGCGCAACATTCATAAGCAGCGCTGATCCGCGCTTGCACAAATACAGCAGGAAG 2270
781 GAAAAAGCGCGCAAGGCGCGCAACGTTGATG 814
2271 GAAAAAGCGCGCAAGGCGCGCAACGTTGATG 2304

RESULT 8

US-08-878-360-7
Sequence 7, Application US/08878360
Patent No. 5945322

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLES OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSER: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/878,360

FILING DATE: 18-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-095B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 5859 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Neisseria gonorrhoeae

STRAIN: F62

FEATURE:

NAME/KEY: CDS

LOCATION: 1491..2330

OTHER INFORMATION: 1gcb

US-08-878-360-7

Query Match 87.7%; Score 726; DB 2; Length 5859;

Best Local Similarity 93.2%; Pred. No. 2.8e-216;

	Matches	759;	Conservative	0;	Mismatches	55;	Indels	0;	Gaps	0;
OY	1	ATGCAAAACCAAGTTATCATAGCTTACCTTCGGCCGGAAGAACCAGAGGGCGCAATTGCCGAT	60							
Db	1491	ATGCAAAAACAAGTTATCAGCTTGACTTCGCCGGAAGAACCAGAGGGCGCAATTGCCGCA	1550							
OY	61	ACCTTCGGCAGGCAAGGCATCCCGTTTCAAGTTTTTGAAGCACTGATCCGTCTGAAGAAG	120							
Db	1551	ACCTTCGGGAGTCGGGGCATCCCGTTTCAAGTTTTTGAAGCACTGATCCGTCTGAAGAAG	1610							
OY	121	CTGGAAACAGGCAATGGCCGAACCTGTCCTCCCGACTTGTGGGGCACCCCTTAATTGACGGA	180							
Db	1611	CTGGAAACAGGCAATGGCCGAACCTGTCCTCCCGACTTGTGGGGCACCCCTTAATTGACGCGA	1670							
OY	181	GTGGAAAAAGCCTGCTTATATGAGCCACGCCGATATGTGAAGACAGGCAATTGAGAGAGT	240							
Db	1671	GTGGAAAAAGCCTGCTTATATGAGCCACGCCGATATGTGAAGACAGGCAATTGAGAGAGGC	1730							
OY	241	CTGCCGTAATCAACCGTAATTTGAGGACAGAGTTTACTCGGCGAAGGTGAAGAAAAATTTC	300							
Db	1731	GTACCGTATATCCGCGTATTTGAAGATGATGCTTACTCGGCGAAGGCGCGAGAGATTTC	1790							
OY	301	CTTGGCCGAAGACGCTTGCTGCAAGAACAGCTTTGAACCCGGAATCCGCTTTATGTCCGC	360							
Db	1791	CTTGGCGAAGATCTTGCTGCAAGAACAGCTTTGAACCCGGAATCCGCTTTGTGTGCTCGC	1850							
OY	361	TTGGAACAGATGTTTATGACAGCTCTGAACCTCGCCCTCGGCGTGGAGATTACTGCGAGG	420							
Db	1851	TTGGAACAGATGTTTATGACAGCTCTGAACCTCGCCCTCGGCGTGGAGATTACTGCGAGGG	1910							
OY	421	CGCGCTTTCCGCTGTGGAAAAAGCAACACTGGGGGAGCGCGGCTATATCAATTTCCTCGA	480							
Db	1911	CGCGCTTTCCGCTGTGGAAAAAGCAACACTGGGGGAGCGCGGCTATATTAATTATTCCTCGA	1970							
OY	481	AAAGGATGCGGTTTCTGGAACAGTTTCCGCGCTCGCCCGCGAAGGGCTGACCCC	540							
Db	1971	AAAGGATGCGGTTTCTGGAACAGTTTCCGCGCTCGCCCGCGAAGGGCTGACCCCC	2030							
OY	541	GTCGATCTGATGATGTTTCAGCGATTTTTCGACAGGGAAAGAAATGCCGTTTGCAGCTC	600							
Db	2031	GTCGATTTGATGATGTTTCAGCGAACCCTGACCAAGGGAAAGAAATGCCGTTTGCAGCTC	2090							
OY	601	AATCCCGGCTTGTGGGCCCAAGAGCTGATTAATGCAAGTTTCAAGACAAACAGAGCGCA	660							
Db	2091	AATCCCGGCTTGTGGGCCCAAGAGCTGATTAATGCAAGTTTCAAGACAAACAGAGCGCA	2150							
OY	661	TTGGGCAAGCTGATCGAAGACGACCGCCTCTCTGAACCGCAACAGCAAAGGCGGATTC	720							
Db	2151	TTGGGCAAGCTGATCGAAGACGACCGCCTCTCTGAACCGCAACAGCAATGCGCGATTC	2210							
OY	721	CCCGCAACCAATTCAAACACCGCTGATCCGCGCTTGAACCAAAATCAGAGGGAAAG	780							
Db	2211	CCCGCAACCAATTCAAACACCGCTGATCCGCGCTTGAACCAAAATGCGAGGAAAG	2270							
OY	781	GA AAAACGCCGCGCAAAAGCGGAAACAGTTTCATTG	814							
Db	2271	GA AAAACGCCGCGCAAAAGCGGCAACAGTTATTCG	2304							
RESULT 9										
US-08-478-140B-1										
; Sequence 1, Application US/08478140B										
; Patent No. 6127153										
; GENERAL INFORMATION:										
; APPLICANT: JOHNSON, KARL F.										
; APPLICANT: ROTH, STEPHEN										
; APPLICANT: BUCZALA, STEPHANIE L.										
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO										
; TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGALACTOSYLTRANSFERASE, A										
; TITLE OF INVENTION: POLYGALACTOSYLTRANSFERASE AND GENE ENCODING A										
; NUMBER OF SEQUENCES: 8										
; CORRESPONDENCE ADDRESS:										

QY 301 CTTCGGAAGACGCTGGCTGCAAGAACGCTTTGACCCGGATACCGGCTTTATGTCGCG 360
DB 1791 CTTCGGAAGATCTTGCTGCAAGAACGCTTTGACCCGGATACCGGCTTTATGTCGCG 1850
QY 361 TTGGAAACGATGTTTATGACAGCTCTGACCTCGGCTCGGCGGGAATTAATGCGGG 420
DB 1851 TTGGAAACGATGTTTATGACAGCTCTGACCTCGGCTCGGCGGGAATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCGCTTGGAAAGCAACATGGGGGCAACCGCGGCTTATATATTTCCGA 480
DB 1911 CGGCGCTTTCGCGCTTGGAAAGCAACATGGGGGCAACCGCGGCTTATATATTTCCGA 1970
QY 481 AAGCGATGCGTTTCTGCAAGGTTGCGCGCTCGCGCGGCAAGGCTGCAACCC 540
DB 1971 AAGCGATGCGTTTCTGCAAGGTTGCGCGCTCGCGCGGCAAGGCTGCAACCC 2030
QY 541 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAGAAATGCGGTTTCCAGCTC 600
DB 2031 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAGAAATGCGGTTTCCAGCTC 2090
QY 601 AATCCGCGCTTGCGCCCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 660
DB 2091 AATCCGCGCTTGCGCCCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 2150
QY 661 TTGGGAGCGCTGATCGAACAGACCGCTCTGCAACCGCAACAGCAAGGCGGATTC 720
DB 2151 TTGGGAGCGCTGATCGAACAGACCGCTCTGCAACCGCAACAGCAAGGCGGATTC 2210
QY 721 CCGGCAACATTCATAACACCGCTGATCGGCGCTTGAACCAAAATCAGCAGGGAAG 780
DB 2211 CCGGCAACATTCATAACACCGCTGATCGGCGCTTGAACCAAAATCAGCAGGGAAG 780
QY 781 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 814
DB 2271 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 2304

RESULT 10
US-09-333-412-1

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-333-412-1

Query Match 87.7%; Score 726; DB 3; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACAGGTTATGAGCTTAGCTTCGCGCGGAAGACGAGGCGGCAATTTGCGAT 60
DB 1491 ATGCAAAACAGGTTATGAGCTTAGCTTCGCGCGGAAGACGAGGCGGCAATTTGCGAT 1550
QY 61 ACCTTGAGGAGGACGCGCAATCCGTTTCAAGTTTTCAGAGCACTGATGCGTCTGAAG 120
DB 1551 ACCTTGAGGAGGACGCGCAATCCGTTTCAAGTTTTCAGAGCACTGATGCGTCTGAAG 1610
QY 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTGCGCGCAACCTTAATTTGACGGA 180
DB 1611 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTGCGCGCAACCTTAATTTGACGGA 1670
QY 181 GTGAAAAAAGCGCTTATGAGCAACGCGGTAATGTGGAAGAGCAATTTGACGAAGGT 240
DB 1671 GTGAAAAAAGCGCTTATGAGCAACGCGGTAATGTGGAAGAGCAATTTGACGAAGGT 1730
QY 241 CTGCGTATATACCGTATTTGAGAGCAAGTTTACTGCGGAAGGTGAGAAAAATTC 300
DB 1731 GTACGATATATCGCGTATTTGAGATATATCTTAATCGGGAAGGCGCGAGCAAGTTC 1790
QY 301 CTTCGGAAGACGCTTGCTGCAAGACGCTTTGACCGGAATACCGCTTTATGCTCGCG 360
DB 1791 CTTCGGAAGATATGCTGCTGCAAGACGCTTTGACCGGAATACCGCTTTATGCTCGCG 1850
QY 361 TTGGAAAGATGTTTATGACAGTCTGACCTGCGCTCGGCGGCGGCAATTAATGCGGG 420
DB 1851 TTGGAAAGATGTTTATGACAGTCTGACCTGCGCTCGGCGGCGGCAATTAATGCGGG 1910
QY 421 CCGCGCTTTCGCTGTGGAAGCAACACTGAGGCGGCACTATATATTTCCGA 480
DB 1911 CCGCGCTTTCGCTGTGGAAGCAACACTGAGGCGGCACTATATATTTCCGA 1970
QY 481 AAGCGATGCGGTTTCTGGAAGGTTTTCGCGCTGCGCGGCAAGGCGTGAACCCC 540
DB 1971 AAGCGATGCGGTTTCTGGAAGGTTTTCGCGCTGCGCGGCAAGGCGTGAACCCC 2030
QY 541 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAGAAATGCGGTTTCCAGCTC 600
DB 2031 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAGAAATGCGGTTTCCAGCTC 2090
QY 601 AATCCGCGCTTGCGCCCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 660
DB 2091 AATCCGCGCTTGCGCCCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 2150
QY 661 TTGGGAGCGCTGATCGAACAGACCGCTCTGAAACGCAACAGCAAGGCGGATTC 720
DB 2151 TTGGGAGCGCTGATCGAACAGACCGCTCTGAAACGCAACAGCAAGGCGGATTC 2210
QY 721 CCGGCAACATTCATAACACCGCTGATCGGCGCTTGAACCAAAATCAGCAGGGAAG 780
DB 2211 CCGGCAACATTCATAACACCGCTGATCGGCGCTTGAACCAAAATCAGCAGGGAAG 2270
QY 781 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 814
DB 2271 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 2304

RESULT 11

US-09-333-412-7
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue

LOCATION: 2342..3262
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3322..4335
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4354..5196
 US-09-338-943-1

Query Match 87.7%; Score 726; DB 3; Length 5859;
 Best Local Similarity 93.2%; Pred. No. 2.8e-216;
 Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGGAGAAAGGAGGAGCAATTGCGCAT 60
 1491 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGGAGAAAGGAGGAGCAATTGCGCA 1550
 61 ACCCTGCGGAGGAGCAAGGCAATCCCGTTTCACTTTTTCAGCTGAGCACTGATGCCGTGAAAG 120
 1551 ACCCTGCGGAGGAGCAATCCCGTTTCACTTTTTCAGCTGAGCACTGATGCCGTGAAAG 1610
 121 CTGGAACAGGCAATGCGGAACTGCTCCCGGCTTGTGCGGCAACCCCTATTGAGCGCA 180
 1611 CTGGAACAGGCAATGCGGAACTGCTCCCGGCTTGTGCGGCAACCCCTATTGAGCGCA 1670
 181 GTGGAAGAAAGCGCTTTATGAGCAAGCGCTATTGGAAGAGGCAATTGAGCAAGAT 240
 1671 GTGGAAGAAAGCGCTTTATGAGCAAGCGCTATTGGAAGAGGCAATTGAGCAAGAT 1730
 241 CTGCGGTATATCAACCGTATTGGAAGAGAGCTTTTATCTCGCGAGAGTGAAGAAATTC 300
 1731 GTACCTATATCGCGGTATTGGAAGAGAGTCTTCTCGCGAGAGGCGAGAGGAGTTC 1790
 301 CTGCGGAGAGAGCTTGTGCGGAGAGAGCTTTGAGACCCGAGATCCGCTTTATGCTCCG 360
 1791 CTGCGGAGAGATCTGTGCGGAGAGAGCTTTGAGACCCGAGATCCGCTTTATGCTCCG 1850
 361 TTGGAAGAGATGTTATGAGCAAGCTGACCTCGGCGCTCGGCGAGGATTAATCACTCGG 420
 1851 TTGGAAGAGATGTTATGAGCAAGCTGACCTCGGCGCTCGGCGAGGATTAATCACTCGG 1910
 421 CGCGCTTTTCCGCTGTGGAAGAGCAATGCGGAGAGCGCGGCTATATCAATTTCCGA 480
 1911 CGCGCTTTTCCGCTGTGGAAGAGCAATGCGGAGAGCGCGGCTATATNTTTCCGA 1970
 481 AAAGGATGCGGTTTCTCTGAGCAGGTTTCCGCTCGCGCGGAGAGGCTGCAACCC 540
 1971 AAAGGATGCGGTTTCTCTGAGCAGGTTTCCGCTCGCGCGGAGAGGCTGCAACCC 2030
 541 GTGCAATCTGATGATGTTGAGCGATTTTTCAGAGGAGAGAAATGCGGTTTCCAGCTC 600
 2031 GTGCAATCTGATGATGTTGAGCGATTTTTCAGAGGAGAGAAATGCGGTTTCCAGCTC 2090
 601 AATCCCGCTTGTGCGGCGCAAGAGCTGATTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
 2091 AATCCCGCTTGTGCGGCGCAAGAGCTGATTAATGCAAGTTTTCAGACCAAAACAGCGCA 2150
 661 TTGGGAGAGCTGATGAGCAAGAGCGGCTTCTGAAACCGCAACAGCAAAAGCGCGATTC 720
 2151 TTGGGAGAGCTGATGAGCAAGAGCGGCTTCTGAAACCGCAACAGCAAAAGCGCGATTC 2210
 721 CCGGCAACAGATTAACAAACAGCGGCTGATCCCGGCTTTCAGCAAAATTCAGAGGAGAG 780
 2211 CCGGCAACAGATTAACAAACAGCGGCTGATCCCGGCTTTCAGCAAAATTCAGAGGAGAG 2270
 781 GAAAAACGCGGCAAGAGCGGCAAGCTTCAATG 814
 2271 GAAAAACGCGGCAAGAGCGGCAAGCTTCAATG 2304

RESULT 13
 US-10-007-267A-1
 ; Sequence 1, Application US/10007267A

Patent No. 6780624
 GENERAL INFORMATION:
 APPLICANT: Gotschlich, Emil C.
 TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
 FILE REFERENCE: 040853-01-5029-02
 CURRENT APPLICATION NUMBER: US/10/007,267A
 CURRENT FILING DATE: 2001-12-03
 PRIOR APPLICATION NUMBER: US 09/333,412
 PRIOR FILING DATE: 1999-06-15
 PRIOR APPLICATION NUMBER: US 08/878,360
 PRIOR FILING DATE: 1997-06-18
 PRIOR APPLICATION NUMBER: US 08/683,426
 PRIOR FILING DATE: 1996-07-18
 PRIOR APPLICATION NUMBER: US 08/312,387
 PRIOR FILING DATE: 1994-09-24
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 1
 LENGTH: 5859
 TYPE: DNA
 ORGANISM: Neisseria gonorrhoeae
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (1)..(381)
 OTHER INFORMATION: 1gfc
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (445)..(1491)
 OTHER INFORMATION: 1gfc
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (2342)..(3262)
 OTHER INFORMATION: 1gfc
 NAME/KEY: misc.feature
 LOCATION: (3322)..(4335)
 OTHER INFORMATION: 1gfc
 NAME/KEY: misc.feature
 LOCATION: (4354)..(5196)
 OTHER INFORMATION: 1gfc
 US-10-007-267A-1

Query Match 87.7%; Score 726; DB 3; Length 5859;
 Best Local Similarity 93.2%; Pred. No. 2.8e-216;
 Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGGAGAAAGGAGGAGCAATTGCGCAT 60
 1491 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGGAGAAAGGAGGAGCAATTGCGCA 1550
 61 ACCCTGCGGAGGAGCAAGGCAATCCCGTTTCACTTTTTCAGCTGAGCACTGATGCCGTGAAAG 120
 1551 ACCCTGCGGAGGAGCAATCCCGTTTCACTTTTTCAGCTGAGCACTGATGCCGTGAAAG 1610
 121 CTGGAACAGGCAATGCGGAACTGCTCCCGGCTTGTGCGGCAACCCCTATTGAGCGCA 180
 1611 CTGGAACAGGCAATGCGGAACTGCTCCCGGCTTGTGCGGCAACCCCTATTGAGCGCA 1670
 181 GTGGAAGAAAGCGCTTTATGAGCAAGCGCTATTGGAAGAGGCAATTGAGCAAGAT 240
 1671 GTGGAAGAAAGCGCTTTATGAGCAAGCGCTATTGGAAGAGGCAATTGAGCAAGAT 1730
 241 CTGCGGTATATCAACCGTATTGGAAGAGAGCTTTTATCTCGCGAGAGTGAAGAAATTC 300
 1731 GTACCTATATCGCGGTATTGGAAGAGAGTCTTCTCGCGAGAGGCGAGAGGAGTTC 1790
 301 CTGCGGAGAGAGCTTGTGCGGAGAGAGCTTTGAGACCCGAGATCCGCTTTATGCTCCG 360
 1791 CTGCGGAGAGATCTGTGCGGAGAGAGCTTTGAGACCCGAGATCCGCTTTATGCTCCG 1850
 361 TTGGAAGAGATGTTATGAGCAAGCTGACCTCGGCGCTCGGCGAGGATTAATCACTCGG 420

Db 1851 TTGAAAACGATGTTATGACAGTCGACCTCCCTCCGCGTGGCGGACTACGCGCGG 1910
QY 421 CCGGCTTTCCGCTGTTGGAAAGCGAACACATGCGGGAACGCGCGGCTATATCATTTCCGA 480
Db 1911 CCGGCTTTCCGCTTTTGGAAAGCGAACACTGCGGGAACGCGCGGCTATATTTATTTCCGA 1970
QY 481 AAGGCGATGCGGTTTTCCTGACAGTTTGGCCCTCGCGCCGGAAGGCGTGCACCC 540
Db 1971 AAGGCGATGCGGTTTTCCTGACAGTTTGGCCCTCGCGCCGGAAGGCGTGCACCC 2030
QY 541 GTGATCTGATGATGTTTCAGCGATTTTTCAGACAGGAAAGAAATGCGGTTTTCAGCTC 600
Db 2031 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACAGGAAAGAAATGCGGTTTTCAGCTC 2090
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATATATGCAAGTTTACAGACCAAAACAGCGA 660
Db 2091 AATCCGCTTGTGCGCCCAAGAGCTGATATATGCAAGTTTACAGACCAAAACAGCGA 2150
QY 661 TTGGGAGCCTGATTCGACACGACCGCTCTGTAACCGCAACAGCAAGCGCGATTC 720
Db 2151 TTGGGAGCCTGATTCGACACGACCGCTCTGTAACCGCAACAGCAATGGCGCATTC 2210
QY 721 CCGGCAACACATTCAAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGCGCAAGCGCGCAACAGTTCAATTG 814
Db 2271 GAAAAACGCGCGCAAGCGCGCAACAGTTAATCG 2304

RESULT 14

US-10-007-267A-7
; Sequence 7, Application US/10007267A
; Patent No. 6780624
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-02
; CURRENT APPLICATION NUMBER: US/10/007,267A
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent# version 3.2
; SEQ ID NO 7
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1491)..(2330)
; OTHER INFORMATION: lgtb
US-10-007-267A-7

Query Match

Best Local Similarity 87.7%; Score 726; DB 3; Length 5859;
Matches /59; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGTATGCTTAGCTTCGCGGACGAGGCGGACATTCGCCAT 60
Db 1491 ATGCAAAACCAAGTATGCTTAGCTTCGCGGACGAGGCGGACATTCGCCA 1550
QY 61 AACTTGGAGGACGCGCATCCCGTTTCAGTTTTCAGACACTGATGCCGTCTGAAG 120

Db 1551 AACTTGGAGGACGCGCATCCCGTTTCAGTTTTCAGACGACTGATGCCGTCTGAAG 1610
QY 121 CTGAAACAGGCAATGGCGAAACTGTTCCCGGCTTGTGCGCACCCCTATTTGACGGA 180
Db 1611 CTGAAACAGGCAATGGCGAAACTGTTCCCGGCTTGTGCGCACCCCTATTTGACGGA 1670
QY 181 GTGAAAAAGCTGCTTTATGACGACCGCGATTTGTGGAAGCAGGCAATTGACGAAGT 240
Db 1671 GTGAAAAAGCTGCTTTATGACGACCGCGATTTGTGGAAGCAGGCAATTGACGAAGT 1730
QY 241 CTGCGTATATCACCCTGATTTTGAAGACAGTTTACTCGGCGAAGTGAAGAAAAATT 300
Db 1731 GTACGTAATATGCGCTATTTGAAGATGATGCTTACTCGGCGAAGGCGGAGCATTC 1790
QY 301 CTGCGGAAAGAGCTTGGCTGCAAAAGCTTTGACCGGATACCGCTTTATGCTCCG 360
Db 1791 GTACGTAATATGCGCTATTTGAAGATGATGCTTACTCGGCGAAGGCGGAGCATTC 1850
QY 361 TTGAAACGATGTTTATGACAGTCTGACCTCGCCTCGGCGTGGCGAATTACTCGCG 420
Db 1851 TTGAAACGATGTTTATGACAGTCTGACCTCGCCTCGGCGTGGCGAATTACTCGCG 1910
QY 421 CCGGCTTTCCGCTGTTGAAAGCGAACACTGCGGGAACGCGGCTATATCATTTCCGA 480
Db 1911 CCGGCTTTCCGCTTTTGAAGCGAACACTGCGGGAACGCGGCTATATTTCCGA 1970
QY 481 AAGGCGATGCGGTTTTCCTGACAGTTTTCGCGCTCGCGCCGGAAGGCTGCACCC 540
Db 1971 AAGGCGATGCGGTTTTCCTGACAGTTTTCGCGCTCGCGCCGGAAGGCTGCACCC 2030
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACAGGAAAGAAATGCGGTTTTCAGCTC 600
Db 2031 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACAGGAAAGAAATGCGGTTTTCAGCTC 2090
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATATATGCAAGTTTACAGACCAAAACAGCGA 660
Db 2091 AATCCGCTTGTGCGCCCAAGAGCTGATATATGCAAGTTTACAGACCAAAACAGCGA 2150
QY 661 TTGGGAGCCTGATTCGACACGACCGCTCTGTAACCGCAACAGCAAGCGCGATTC 720
Db 2151 TTGGGAGCCTGATTCGACACGACCGCTCTGTAACCGCAACAGCAATGGCGCATTC 2210
QY 721 CCGGCAACACATTCAAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGCGCAAGCGCGCAACAGTTCAATTG 814
Db 2271 GAAAAACGCGCGCAAGCGCGCAACAGTTAATCG 2304

RESULT 15

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text

```

1      CURRENT APPLICATION DATA:
2      APPLICATION NUMBER: US/09/557,884
3      FILING DATE: 25-Apr-2000
4      CLASSIFICATION: <Unknown>
5
6      PRIOR APPLICATION DATA:
7      APPLICATION NUMBER: 08/476,102
8      FILING DATE: JUN-5-1995
9
10     ATTORNEY/AGENT INFORMATION:
11
12     NAME: Michelle S. Marks
13     REGISTRATION NUMBER: 41,971
14     REFERENCE/DOCKET NUMBER: PB186P3
15
16     TELECOMMUNICATION INFORMATION:
17     TELEPHONE: 301-309-8504
18     TELEFAX: 301-309-8439
19
20     INFORMATION FOR SEQ ID NO: 1:
21
22     SEQUENCE CHARACTERISTICS:
23
24     LENGTH: 1830121 base pairs
25     TYPE: nucleic acid
26     STRANDEDNESS: double
27     TOPOLOGY: linear
28
29     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
30
31     US-09-557-884-1

```

Query Match	10.5%;	Score 86.6;	DB 3;	Length 1830221;
Best Local Similarity	51.0%;	Pred. No. 3.6e-15;		
Matches 231; Conservative	0;	Mismatches 219;	Indels 3;	Gaps 1;

Qy	180	GGTGAAGAAAAGCGTCGTTATATAGACGACCGCCGATTTGAGAAAGCAGGACANTGGACGAAGG	239
Db	570747	AGGAGAAAAGCGCTGTTATATATAGCCATTTCTTATATATGAAATATAGTGTGATGATGAAAA	5706888
Qy	240	TCGCGCGTATATACCCGTAATTTTGAGGACGACGTTTTACTCGCGGAAGTGAGGAAAAATT	299
Db	570687	TTTTGAATATCTCAAAATTTTTGAGATGATGTAATTTCTGGGGAAGAAATGCGGAATATTT	570628
Qy	300	CCTTGCCGGAAGACGCTTGCGCTGCAGAGAAGGCTTTGACCCGGATACCGCCTTTATCGTCCG	359
Db	570637	TTTGAACCAAAATGAAATGGTATAAAAACAGGTTTGTATTTAATGATATTTTATATTTGG	570568
Qy	360	CTTGGAAAAGATGTTATATGACAGT---CCTGACCTTGCCCTCGCGGTCGAGCTACTAG	416
Db	570567	TTTAAAAACTTTTTTAAACGACGATTAAACTTTGAGAAACAAACTTAAATTTCCACCTTTAA	570508
Qy	417	CGGCGCGCGCTTTCCGCTGTGGAAAAGCGACACTGCGGGGAGCGCGCGCTATATCATTTTC	476
Db	570507	TTCTTAGGAACCTTGAATATTTTAAATTCACACTCGGGGGGAGCGCGGTTATATATTTTC	570448
Qy	477	CGGAAAAGGATGCGCGTTTTTCTGTGACAGGTTTTCGCGCCCTGCGCGCGGAAAGCGCTGCA	536
Db	570447	TCAAGGTGCGGCTAATATATGTAATGTAATGTAATTTAAAGAAATATTTCTTCTGATGAAATTTGT	570388
Qy	537	CCCGCGTCAATCTGATGATGATGTCAGCGAATTTTTTCGACAGGGAAGAAATGCGGTTTTGCCA	596
Db	570387	TGCAATTTGATGAACCTAATTTTTTAATAAATTAAGTGTGATGTTGAATTAATATATGTCATACA	570328
Qy	597	GCTCAATCCCGCTTGTGCGCCCAAGAGCTGCA	629
Db	570327	ACTGATCCAGCAATTTGTATTTCAAGAACTTCCA	570295

Search completed: April 7, 2006, 18:22:52
Job time : 305.232 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:02:11 ; Search time 1357.05 Seconds
(without alignments)
5045.535 Million cell updates/sec

Title: US-09-211-691-1

Sequence: 1 atgcaaacaccgttatcag.....tcattgtgccttccaataa 828

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database : Published Applications NA_Main:*
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Decryption
1	828	100.0	828	3	US-09-211-691-1	Sequence 1, Appl1
2	828	100.0	828	6	US-10-317-773-1	Sequence 1, Appl1
3	828	100.0	828	6	US-10-317-7428-1	Sequence 1, Appl1
C 4	826.4	99.8	4659.4	9	US-10-915-740A-4	Sequence 4, Appl1
C 5	826.4	99.8	22421.6	9	US-10-915-740A-1068	Sequence 1068, Appl1
6	726	87.7	585.9	5	US-10-007-267-1	GENERAL INFORMA
7	726	87.7	585.9	5	US-10-007-267-7	GENERAL INFORMA
8	726	87.7	585.9	6	US-10-096-125-1	Sequence 1, Appl1
9	726	87.7	585.9	7	US-10-654-528-1	Sequence 1, Appl1
10	726	87.7	585.9	7	US-10-654-528-7	Sequence 7, Appl1
11	426	51.4	840	8	US-10-472-260-175	Sequence 175, Appl
12	93	11.2	1454.9	9	US-10-795-155-681	Sequence 570, Appl
13	93	11.2	1066.5	9	US-10-795-155-681	Sequence 681, Appl
C 14	86.6	10.5	18301.21	7	US-10-329-670-1	Sequence 1, Appl1
C 15	86.6	10.5	18301.21	8	US-10-158-865-1	Sequence 1, Appl1
C 16	86.6	10.5	18301.21	9	US-10-981-667-1	Sequence 1, Appl1
C 17	73.2	8.8	458.3	9	US-10-844-776-16	Sequence 16, Appl1
C 18	73.2	8.8	558.4	9	US-10-844-776-17	Sequence 17, Appl1
C 19	73.2	8.8	6579.2	7	US-10-672-787-31	Sequence 31, Appl1
C 20	71.4	8.6	5570	9	US-10-844-776-15	Sequence 15, Appl1
21	57	6.9	3625	9	US-10-795-159-517	Sequence 517, Appl
22	57	6.9	13335.6	9	US-10-795-159-682	Sequence 682, Appl
23	57	6.9	18301.21	7	US-10-329-670-1	Sequence 1, Appl1

C	25	57	6.9	1830121	8	US-10-158-865-1	Sequence 1, Appl1
C	24	57	6.9	1830121	9	US-10-981-687-1	Sequence 1, Appl1
C	26	40.2	4.9	1351	7	US-10-767-701-13469	Sequence 13469, A
C	27	39.8	4.8	4332	10	US-11-097-143-13511	Sequence 13511, A
C	28	39	4.7	984	10	US-11-097-143-34813	Sequence 34813, A
C	29	39	4.7	3068	10	US-11-097-143-34813	Sequence 34813, A
C	30	38.6	4.7	954	6	US-10-369-493-40566	Sequence 40566, A
C	31	37.8	4.6	2511	7	US-10-451-866-2	Sequence 2, Appl1
C	32	37.8	4.6	4540	7	US-10-451-866-4	Sequence 4, Appl1
C	33	37.4	4.5	1329	8	US-10-739-930-5413	Sequence 5413, Ap
C	34	37.2	4.5	760	5	US-10-027-632-7118	Sequence 7118, Ap
C	35	37.2	4.5	760	5	US-10-027-632-7119	Sequence 7119, Ap
C	36	37.2	4.5	760	6	US-10-027-632-7118	Sequence 7118, Ap
C	37	37.2	4.5	760	6	US-10-027-632-7119	Sequence 7119, Ap
C	38	36.8	4.4	14643	6	US-10-156-761-931	Sequence 931, App
C	39	36.8	4.4	31422	7	US-10-204-862A-2	Sequence 2, Appl1
C	40	36.8	4.4	31422	10	US-11-005-196-2	Sequence 2, Appl1
C	41	36.8	4.4	9025608	6	US-10-156-761-931	Sequence 3048, A
C	42	36.4	4.4	1122	7	US-10-287-122A-32048	Sequence 32048, A
C	43	36.4	4.4	7788	6	US-10-329-079-8	Sequence 8, Appl1
C	44	36.4	4.4	12911	9	US-10-915-740A-54	Sequence 54, Appl1
C	45	36.4	4.4	37360	6	US-10-329-079-6	Sequence 6, Appl1

ALIGNMENTS

```

RESULT 1
US-09-211-691-1
Sequence 1, Application US/09211691
Patent No. US20020034805A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
APPLICANT: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 019957-012910US
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 828
TYPE: DNA
ORGANISM: Neisseria meningitidis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(828)
OTHER INFORMATION: beta-1,4-galactosyltransferase (1gbt)
US-09-211-691-1

```

Query Match	100.0%;	Score 828;	DB 3;	length 828;
Best Local Similarity	100.0%;	Pred. No. 6.6e-267;		
Matches 828; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	1	ATGCAAAACCAAGCTTATCAGCTTAGCTTCCCGCGAGAAACGAGGGCGCAATGCGCGAT	60
Db	1	ATGCAAAACCAAGCTTATCAGCTTAGCTTCCCGCGAGAAACGAGGGCGCAATGCGCGAT	60
QY	61	ACCTTGGCAGGCAAGCGCATCCGTTTGAATTTTCAGACGACCTGATGCGGTCTGAAGG	120
Db	61	ACCTTGGCAGGCAAGCGCATCCGTTTGAATTTTCAGACGACCTGATGCGGTCTGAAGG	120
QY	121	CTGGAACAGGCAATGCGCGAATCTGTCCTCCCGGCTTGTTCGCGACCCCTCATTTTGAAGCGGA	180
Db	121	CTGGAACAGGCAATGCGCGAATCTGTCCTCCCGGCTTGTTCGCGACCCCTCATTTTGAAGCGGA	180
QY	181	GTGAAAAAGCCTGCTTTATGAGCCAGCGCGTATTTGTGGAAGCAGGCAATGAGCAAGGT	240

Db 181 GTGAAAAAGCGCTTTATGAGCAGCGCGTATTGTGAGAGAGGCAATTGAGCAAGGT 240
QY 241 CTGCGGTATATCAACCGTATTTGAGAGAGAGGCTTTTACTCGGCGAAGGTGAGAAAAATTC 300
Db 241 CTGCGGTATATCAACCGTATTTGAGAGAGAGGCTTTTACTCGGCGAAGGTGAGAAAAATTC 300
QY 301 CTGCGAAGAGCGCTTGAGTGAAGAGCGCTTGACCCGGAATACCGCTTTATCGTCCG 360
Db 301 CTGCGAAGAGCGCTTGAGTGAAGAGCGCTTGACCCGGAATACCGCTTTATCGTCCG 360
QY 361 TTGGAAGACATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGATTATCGCG 420
Db 361 TTGGAAGACATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGATTATCGCG 420
QY 421 CGCGCTTTCCGCTGTGGAAGAGCACTGAGGAGAGCGCGGCTATATCATTTCCGA 480
Db 421 CGCGCTTTCCGCTGTGGAAGAGCACTGAGGAGAGCGCGGCTATATCATTTCCGA 480
QY 481 AAGCGATGCGGTTTTCCTGAGCAGGTTTGCCTGCGCCGCGAGAGGCTGACCCG 540
Db 481 AAGCGATGCGGTTTTCCTGAGCAGGTTTGCCTGCGCCGCGAGAGGCTGACCCG 540
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAGAAATGCGGTTTGCAGCTC 600
Db 541 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAGAAATGCGGTTTGCAGCTC 600
QY 601 AATCCCGCTTGTGCGCCGAGAGGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 601 AATCCCGCTTGTGCGCCGAGAGGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGAGGAGCGCTGATTCGACAGCAGCGCTCTGAGACCGGAAACAGGAGGCGGATTC 720
Db 661 TTGAGGAGCGCTGATTCGACAGCAGCGCTCTGAGACCGGAAACAGGAGGCGGATTC 720
QY 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTTCGAAATTCAGCGGAGAAAG 780
Db 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTTCGAAATTCAGCGGAGAAAG 780
QY 781 GAAAAAGCGCGGCAAGGCGGCAAGTTCATTGTGCTTTCCAAATTA 828
Db 781 GAAAAAGCGCGGCAAGGCGGCAAGTTCATTGTGCTTTCCAAATTA 828

RESULT 2
US-10-317-773-1
; Sequence 1, Application US/10317773
; Publication No. US20030180928A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNAc 4' Epimerase and a
; FILE REFERENCE: 019633-000812US
; CURRENT APPLICATION NUMBER: US/10/317,773
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(828)
; OTHER INFORMATION: beta-1,4-galactosyltransferase (1gcb)
US-10-317-773-1

Query Match 100.0%; Score 828; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 6,66-267;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGCTTATGAGCTTACCTCCGCGGAGAAACGAGGCGCCACATTCGGAT 60
Db 1 ATGCAAAACCAAGCTTATGAGCTTACCTCCGCGGAGAAACGAGGCGCCACATTCGGAT 60
QY 61 ACCCTGAGGAGCAGCGGATCCGTTTCAGTTTTCAGAGCACTGATGCGCTGAAAAG 120
Db 61 ACCCTGAGGAGCAGCGGATCCGTTTCAGTTTTCAGAGCACTGATGCGCTGAAAAG 120
QY 121 CTGGAACAGGCAATGAGCGGAACTCGTCCCGGCTGTGAGGCAACCGCTATTTGAGCGGA 180
Db 121 CTGGAACAGGCAATGAGCGGAACTCGTCCCGGCTGTGAGGCAACCGCTATTTGAGCGGA 180
QY 181 GTGAAAAAGCGCTGTTTATGAGCAGCGCGTATTTGGAAGAGCAATTGAGAGAGT 240
Db 181 GTGAAAAAGCGCTGTTTATGAGCAGCGCGTATTTGGAAGAGCAATTGAGAGAGT 240
QY 241 CTGCGGTATATCAACCGTATTTGAGGAGAGCTTTTACCGGAGAGGAGGAAAAATTC 300
Db 241 CTGCGGTATATCAACCGTATTTGAGGAGAGCTTTTACCGGAGAGGAGGAAAAATTC 300
QY 301 CTGCGCAAGAGCGTTGCTGCAAGAGCGTTTTCAGCCGGAATCCGCTTTATGCTCCG 360
Db 301 CTGCGCAAGAGCGTTGCTGCAAGAGCGTTTTCAGCCGGAATCCGCTTTATGCTCCG 360
QY 361 TTGGAAGACATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGGATTAATCGGAG 420
Db 361 TTGGAAGACATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGGATTAATCGGAG 420
QY 421 CGCGCTTTCCGCTGTTGAAAAGAGCAACATGAGGAGAGCGGCGTATATCATTTCCGA 480
Db 421 CGCGCTTTCCGCTGTTGAAAAGAGCAACATGAGGAGAGCGGCGTATATCATTTCCGA 480
QY 481 AAGGATGCGGTTTTCCTGACAGGTTTCCGCGCTGCGCCGAGAGGCGTCAACCC 540
Db 481 AAGGATGCGGTTTTCCTGACAGGTTTCCGCGCTGCGCCGAGAGGCGTCAACCC 540
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAGAAATGCGGTTTGCAGCTC 600
Db 541 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAGAAATGCGGTTTGCAGCTC 600
QY 601 AATCCCGCTTGTGCGCCGAGAGGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 601 AATCCCGCTTGTGCGCCGAGAGGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGAGGAGCGCTGATTCGACAGCAGCGCTCTGAGACCGGAAACAGGAGGCGGATTC 720
Db 661 TTGAGGAGCGCTGATTCGACAGCAGCGCTCTGAGACCGGAAACAGGAGGCGGATTC 720
QY 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTTCGAAATTCAGCGGAGAAAG 780
Db 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTTCGAAATTCAGCGGAGAAAG 780
QY 781 GAAAAAGCGCGGCAAGGCGGCAAGTTCATTGTGCTTTCCAAATTA 828
Db 781 GAAAAAGCGCGGCAAGGCGGCAAGTTCATTGTGCTTTCCAAATTA 828

RESULT 3
US-10-317-428-1
; Sequence 1, Application US/10317428
; Publication No. US20030186414A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
; FILE REFERENCE: 019633-00081US
; CURRENT APPLICATION NUMBER: US/10/317,428

```

? CURRENT FILING DATE: 2002-12-11
? PRIOR APPLICATION NUMBER: US 60/069,443
? PRIOR FILING DATE: 1997-12-15
? PRIOR APPLICATION NUMBER: US 09/211,691
? PRIOR FILING DATE: 1998-12-14
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1
? LENGTH: 828
? TYPE: DNA
? ORGANISM: Neisseria meningitidis
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(828)
? OTHER INFORMATION: beta-1,4-galactosyltransferase [gcb]
US-10-317-428-1

```

Query Match	100.0%	Score 828	DB 6	Length 828
Best Local Similarity	100.0%	Pred. No. 6	6e-267	
Matches 828; Conservative	0	Mismatches	0	Gaps 0

QY	ATGCAAAACACCGTATACAGCTTAGCTTCGCGCAAAACGAGGGCGCAATTCGCCAT	60
Dp	1	
QY	ACCTTCGGCAGGACGGCAATCCCGTTTCACTTTTTCAGACGCTGATCCGCTTGAAAG	120
Dp	61	
QY	ACCTTCGGCAGGACGGCAATCCCGTTTCACTTTTTCAGACGCTGATCCGCTTGAAAG	120
Dp	121	
QY	CTGGAAACAGGCANATGGCGGAACCTCGCCCGGCTGTGGGCAACCCCTATTGAGGGA	180
Dp	121	
QY	CTGGAAACAGGCANATGGCGGAACCTCGCCCGGCTGTGGGCAACCCCTATTGAGGGA	180
Dp	121	
QY	GTGGAAAAAGCCTGCTTTATAGACCCAGCGATTTGTGAAAGCAGGCAATTGACGAAGT	240
Dp	181	
QY	GTGGAAAAAGCCTGCTTTATAGACCCAGCGATTTGTGAAAGCAGGCAATTGACGAAGT	240
Dp	181	
QY	CTGCGGTATATCACCGTATTTTGAAGACGAGTTTACTCGGCGAAGGTGAGAAAAATTC	300
Dp	241	
QY	CTGCGGTATATCACCGTATTTTGAAGACGAGTTTACTCGGCGAAGGTGAGAAAAATTC	300
Dp	241	
QY	CTTGCCGGAAGACGCTTGGCTGCAGAAAGCCTTTGACCCGGATACGCGCTTATATGTCCG	360
Dp	301	
QY	CTTGCCGGAAGACGCTTGGCTGCAGAAAGCCTTTGACCCGGATACGCGCTTATATGTCCG	360
Dp	301	
QY	TTGGAACAAGATGTTTATAGACGCTCGACCTCGCCCTCGGCGTATATCACTTTCGGA	420
Dp	361	
QY	TTGGAACAAGATGTTTATAGACGCTCGACCTCGCCCTCGGCGTATATCACTTTCGGA	420
Dp	361	
QY	CGCGCTTTCGCGTGTGGAAAGCGAACACTGGGGGACGCGGGCTATATCATTTCCGA	480
Dp	421	
QY	CGCGCTTTCGCGTGTGGAAAGCGAACACTGGGGGACGCGGGCTATATCATTTCCGA	480
Dp	421	
QY	CGCGCTTTCGCGTGTGGAAAGCGAACACTGGGGGACGCGGGCTATATCATTTCCGA	480
Dp	421	
QY	AAAGCGATGGGTTTCTCTGCAACAGGTTTGCGCCCTGCGCCCGAAGGGCTGACCCC	540
Dp	481	
QY	AAAGCGATGGGTTTCTCTGCAACAGGTTTGCGCCCTGCGCCCGAAGGGCTGACCCC	540
Dp	481	
QY	GTGATCTGATGATGTTTCAGCGATTTTTTTCGACAGGGAAGAAATGCCGTTTGCACATC	600
Dp	541	
QY	GTGATCTGATGATGTTTCAGCGATTTTTTTCGACAGGGAAGAAATGCCGTTTGCACATC	600
Dp	541	
QY	AATCCGCGCTTGTGCACCAGAGGCTGATATATGCAAGTTTCAAGACCAAAACAGGCA	660
Dp	601	
QY	AATCCGCGCTTGTGCACCAGAGGCTGATATATGCAAGTTTCAAGACCAAAACAGGCA	660
Dp	601	
QY	TTGGGCAAGCTGATGCAACAGCACGCGCTCTGAAACGCAAAACAGCAAGCTTCC	720
Dp	661	
QY	TTGGGCAAGCTGATGCAACAGCACGCGCTCTGAAACGCAAAACAGCAAGCTTCC	720
Dp	661	
QY	CCGGCAACCAATTCAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG	780
Dp	721	
QY	CCGGCAACCAATTCAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG	780
Dp	721	

[illegible]

RESULT 4
US-10-915-740A-4/C
; Sequence 4, Application US/10915740A

```

, GENERAL INFORMATION:
, APPLICANT: Frazer, Claire M.
, APPLICANT: Hickey, Erin
, APPLICANT: Peterson, Jeremy
, APPLICANT: Tettelin, Hervé
, APPLICANT: Venter, J. Craig
, APPLICANT: Maignani, Vega
, APPLICANT: Galeotti, Cesira
, APPLICANT: Mora, Manroosa
, APPLICANT: Ratti, Giulio
, APPLICANT: Scarselli, Maria
, APPLICANT: Scariato, Vincenzo
, APPLICANT: Rappuoli, Rino
, APPLICANT: Piza, Mariagretta
, APPLICANT: Grandi, Guido
, TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use

```

Query Match	99.8%	Score 826.4	DB 9	Length 46594
Best Local Similarity	99.9%	Pred. No. 1.1e-265		
Matches 827	Conservative 0	Mismatches 1	Indels 0	Gaps 0

Qy	1	ATGCAAAACCA	CGTTATACG	CTTAGCTTCCG	CGCGAAG	CGAGGGCG	CAATTCGCAT	60
Db	33537	ATGCAAAACCA	CGTTATACG	CTTAGCTTCCG	CGCGAAG	CGAGGGCG	CAATTCGCAT	33478
Qy	61	ACCTTGCGC	GAGGCA	CGGCATCCG	GTTCAGTTTTC	GACGCACTG	ATGCGCTCGA	120
Db	33477	ACCTTGCGC	GAGGCA	CGGCATCCG	GTTCAGTTTTC	GACGCACTG	ATGCGCTCGA	33418
Qy	121	CTGGAACAGG	CAATGCG	GAACTCGTCCG	CGCTTGCGG	CGACACCCT	TATTTGAGCGGA	180
Db	33417	CTGGAACAGG	CAATGCG	GAACTCGTCCG	CGCTTGCGG	CGACACCCT	TATTTGAGCGGA	33358
Qy	181	GTCGAAAAAG	CCCTGCTT	TATGAGCC	AGCGCGAT	TGTGGAAC	AGGCAATTTGA	240
Db	33357	GTCGAAAAAG	CCCTGCTT	TATGAGCC	AGCGCGAT	TGTGGAAC	AGGCAATTTGA	33298
Qy	241	CTGCGCGT	ATATCA	CGTATTTT	GAGGACG	AGTTTTA	CTCGGGAAG	300
Db	33297	CTGCGCGT	ATATCA	CGTATTTT	GAGGACG	AGTTTTA	CTCGGGAAG	33238
Qy	301	CTTGCGGA	AGACGCT	TGCGTGA	AGAACCTT	TGACCGGAT	ACCGCTTTAT	360
Db	33237	CTTGCGGA	AGACGCT	TGCGTGA	AGAACCTT	TGACCGGAT	ACCGCTTTAT	33178
Qy	361	TTGGAACG	ATGTTTAT	GACGCTCT	GACCTTCG	CGCGTGA	CGGATTA	420

```
Db 33117 TTGGAACGATGTTTATGACGTCCTGACCTCCGCGCGCTGCGGATTACTCGGG 33118
Qy 421 CGGCGCTTCCGCTGTTGAAAGCAACACTGAGGAGACGCGCGGCTATATCATTTCCGA 480
Db 33117 CGGCGCTTCCGCTGTTGAAAGCAACACTGAGGAGACGCGCGGCTATATCATTTCCGA 33058
Qy 481 AAAGCAGTGGGTTTCTCTGACAGGTTTCCGCGCTGCGCGCAAGGCTGCAAGG 540
Db 33057 AAAGCAGTGGGTTTCTCTGACAGGTTTCCGCGCTGCGCGCAAGGCTGCAAGG 32998
Qy 541 GTGATCTGATGATGTTTACGCGATTTTTCGACAGGAGAAATGCGGTTTCCGCGCTC 600
Db 32997 GTGATCTGATGATGTTTACGCGATTTTTCGACAGGAGAAATGCGGTTTCCGCGCTC 32938
Qy 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAGCAACCAAGCGCA 660
Db 32937 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAGCAACCAAGCGCA 32878
Qy 661 TTGGGAGCGCTGATGCAACGACCGCGCTTCTGTAACCGCAACAGCAAGGCGCATTC 720
Db 32877 TTGGGAGCGCTGATGCAACGACCGCGCTTCTGTAACCGCAACAGCAAGGCGCATTC 32818
Qy 721 CCGGCAACACATTTCAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db 32817 CCGGCAACACATTTCAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 32758
Qy 781 GAAAAAGCGCGGCAAGGCGCGCAAGCTTCAATGTCCTTTCCAAATTA 828
Db 32757 GAAAAAGCGCGGCAAGGCGCGCAAGCTTCAATGTCCTTTCCAAATTA 32710
```

```
RESULT 5
US-10-915-740A-1068/C
/ Sequence 1068, Application US/10915740A
/ Publication No. US20050191316A1
/ GENERAL INFORMATION:
/ APPLICANT: Frazer, Claire M.
/ APPLICANT: Hickey, Brian
/ APPLICANT: Peterson, Jeremy
/ APPLICANT: Tetrelin, Hervé
/ APPLICANT: Venter, J. Craig
/ APPLICANT: Masignani, Vega
/ APPLICANT: Galeotti, Cesira
/ APPLICANT: Mora, Manroba
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Scarselli, Maria
/ APPLICANT: Scariato, Vincenzo
/ APPLICANT: Rappuoli, Rino
/ APPLICANT: Pizzi, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
/ FILE REFERENCE: 002441.00090
/ CURRENT APPLICATION NUMBER: US/10/915,740A
/ PRIOR FILING DATE: 2004-08-11
/ PRIOR APPLICATION NUMBER: 09/806,866
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: USSN 60/103,794
/ PRIOR FILING DATE: 1998-10-09
/ PRIOR APPLICATION NUMBER: USSN 60/132,068
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/US99/25373
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 1068
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1068
/ LENGTH: 2242716
/ TYPE: DNA
/ ORGANISM: Neisseria meningitidis
US-10-915-740A-1068
```

Query Match 99.8%; Score 826.4; DB 9; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 4,6e-265;

```
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGCAAAACGAGTTATCAGCTTACCTTCCGCGGAGAAACGAGGCGCAATTCGGAT 60
Db 1997047 ATGCAAAACGAGTTATCAGCTTACCTTCCGCGGAGAAACGAGGCGCAATTCGGAT 1996988
Qy 61 ACCTTCGCGAGGCAAGGATCCGTTTCTGATGTTTTCAGAGCACTGATGTCGCTGAAAG 120
Db 1996987 ACCTTCGCGAGGCAAGGATCCGTTTCTGATGTTTTCAGAGCACTGATGTCGCTGAAAG 1996928
Qy 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTGCGGACCCCTTATTTGAGCGGA 180
Db 1996927 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTGCGGACCCCTTATTTGAGCGGA 1996868
Qy 181 GTGAAAAAGCGCTGTTATGAGCAACGCGCTATTTGGAAGCAGGCAATGGAAGAGT 240
Db 1996867 GTGAAAAAGCGCTGTTATGAGCAACGCGCTATTTGGAAGCAGGCAATGGAAGAGT 1996808
Qy 241 CTGCGGTATATACCGTATTTGAGAGCAGCTTTTACTCGCGCAAGGTGAGGAAATTC 300
Db 1996807 CTGCGGTATATACCGTATTTGAGAGCAGCTTTTACTCGCGCAAGGTGAGGAAATTC 1996748
Qy 301 CTGCGCAAGAGCGCTTGTGCTGCAAGAACGTTTGAACCGGATACCGCTTATGCTCCG 360
Db 1996747 CTGCGCAAGAGCGCTTGTGCTGCAAGAACGTTTGAACCGGATACCGCTTATGCTCCG 1996688
Qy 361 TTGGAACGATGTTTATGCAAGCTGCTGACCTTCCGCGCTTCCGCGATTAATGCTCGGG 420
Db 1996687 TTGGAACGATGTTTATGCAAGCTGCTGACCTTCCGCGCTTCCGCGATTAATGCTCGGG 1996628
Qy 421 CGGCGCTTCCGCGGTTGGAAGGAAACACTGAGGAGACGCGGCGCTATATCATTTCCGA 480
Db 1996627 CGGCGCTTCCGCGGTTGGAAGGAAACACTGAGGAGACGCGGCGCTATATCATTTCCGA 1996568
Qy 481 AAAGGATGCGGTTTCTCTGACAGGTTTCCGCGCTGCGCGCAAGGCTGCAAGG 540
Db 1996567 AAAGGATGCGGTTTCTCTGACAGGTTTCCGCGCTGCGCGCAAGGCTGCAAGG 1996508
Qy 541 GTGATCTGATGATGTTTCAAGGATTTTTCAGACAGGAGAAATGCGGTTTCCGAGCTC 600
Db 1996507 GTGATCTGATGATGTTTCAAGGATTTTTCAGACAGGAGAAATGCGGTTTCCGAGCTC 1996448
Qy 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAGCAACCAAGCGCA 660
Db 1996447 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAGCAACCAAGCGCA 1996388
Qy 661 TTGGGAGCGCTGATGCAACGACCGCGCTTCTGAAACCGCAACAGCAAGGCGCATTC 720
Db 1996387 TTGGGAGCGCTGATGCAACGACCGCGCTTCTGAAACCGCAACAGCAAGGCGCATTC 1996328
Qy 721 CCGGCAACACATTTCAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db 1996327 CCGGCAACACATTTCAACACCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 1996268
Qy 781 GAAAAAGCGCGGCAAGGCGCGCAAGCTTCAATGTCCTTTCCAAATTA 828
Db 1996267 GAAAAAGCGCGGCAAGGCGCGCAAGCTTCAATGTCCTTTCCAAATTA 1996220
```

```
RESULT 6
US-10-007-267-1
/ GENERAL INFORMATION:
/ APPLICANT: Gotschlich, Emil C.
/ TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
/ OLIGOSACCHARIDES, AND GENES ENCODING THEM
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Klauber & Jackson
/ STREET: 411 Hackensack Avenue
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
```


COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-007-267-1

Query Match 87.7%; Score 726; DB 5; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.5e-232;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTATGAGCTTACGCTTACGCGCGCAAGACGAGGCGCATTTGCCGAT 60
DB 1491 ATGCAAAACCAAGTATGAGCTTACGCTTACGCGCGCAAGACGAGGCGCATTTGCCGCA 1550
QY 61 ACCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 1551 ACCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1610
QY 121 CTGGAACAGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 1611 CTGGAACAGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1670
QY 181 GTGGAAGAAAGCTGCTTATGAGCAAGCGGATTTGTGAAGAGGAGGAGGAGGAGGAGGAGG 240
DB 1671 GTGGAAGAAAGCTGCTTATGAGCAAGCGGATTTGTGAAGAGGAGGAGGAGGAGGAGGAGG 1730
QY 241 CTGCGGTATATACCGTATTTGAGAGCAAGCTTTTACTCGGCGAAGGAGGAGGAGGAGGAGG 300
DB 1731 GTACCGTATATACCGTATTTGAGAGCAAGCTTTTACTCGGCGAAGGAGGAGGAGGAGGAGG 1790
QY 301 CTGCGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 1791 CTGCGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1850
QY 361 TTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 1851 TTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1910
QY 421 CGGCGCTTTCGCTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 1911 CGGCGCTTTCGCTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1970
QY 481 AAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 1971 AAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2030
QY 541 GTGCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 2031 GTGCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2090
QY 601 AATCCCGGCTTGGCGGCGCAAGAGCTGATTTATGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 2091 AATCCCGGCTTGGCGGCGCAAGAGCTGATTTATGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 2150

QY 661 TTGGGAGGCTGATGCAACGAGCGGCTCTGAAACCGCAAAAGGAGGAGGAGGAGGAGGAGG 720
DB 2151 TTGGGAGGCTGATGCAACGAGCGGCTCTGAAACCGCAAAAGGAGGAGGAGGAGGAGGAGG 2210
QY 721 CCGGCAACATTCATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 2211 CCGGCAACATTCATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2270
QY 781 GAAAAAGCGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 2271 GAAAAAGCGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2304

RESULT 7

US-10-007-267-7

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSER: Klauber & Jackson

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,267

FILING DATE: 03-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/333,412

FILING DATE: 15-Jun-1999

APPLICATION NUMBER: 08/312,387

FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-095

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-007-267-7

Query Match 87.7%; Score 726; DB 5; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.5e-232;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTATGAGCTTACGCTTACGCGCGCAAGACGAGGAGGAGGAGGAGGAGGAGG 60
DB 1491 ATGCAAAACCAAGTATGAGCTTACGCTTACGCGCGCAAGACGAGGAGGAGGAGGAGGAGGAGG 1550
QY 61 ACCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 1551 ACCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1610
QY 121 CTGGAACAGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 1611 CTGGAACAGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1670
QY 181 GTGGAAGAAAGCTGCTTATGAGCAAGCGGATTTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 1671 GTGGAAGAAAGCTGCTTATGAGCAAGCGGATTTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGG 1730
QY 241 CTGCGGTATATACCGTATTTGAGAGCAAGCTTTTACTCGGCGAAGGAGGAGGAGGAGGAGGAGG 300

Db 1731 GTACCTTATATGCGCTGATTTGAAAGATGATGTTTCTGCGCAAGGCGCGAGCACTTC 1790
Qy 301 CTTGCCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGGAATACGCTTTATTCGTCGC 360
Db 1791 CTTGCCGAAGATCTTGGCTGCAAGAACGCTTTGACCCGGAATTCGCTTTGTCGTCGC 1850
Qy 361 TTGGAACAGATGTTTATGACAGCTGACCTGACCTCCGCGGTGCGGATTAATGCGG 420
Db 1851 TTGGAACAGATGTTTATGACAGCTGACCTGACCTCCGCGGTGCGGATTAATGCGG 1910
Qy 421 CGGCGCTTTCCGCTGTGGAAGCGAACCTGGGAGACGCGGCTATATCAATTTCCGA 480
Db 1911 CGGCGCTTTCCGCTTTTGGAAAGCGAACCTGGGAGACGCGGCTATATTTATTTCCGA 1970
Qy 481 AAGGCGATGCGGTTTTCCTGACAGGTTTGGCGCCCTCCGCGCAAGGCTGCACTCC 540
Db 1971 AAGGCGATGCGGTTTTCCTGACAGGTTTGGCGCCCTCCGCGCAAGGCTGCACTCC 2030
Qy 541 GTGCAATCTGATGATGTTTCAAGCGATTTTTCGACAGGGAAGATGCGGTTTGGCAGCTC 600
Db 2031 GTGCAATTTATGATGTTTTCGCAACCTGACAGAGGAAGATGCGGTTTGGCAGCTC 2090
Qy 601 AATCCCGCTTGGCGCCCAAGAGCTGATTCGTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCCGCTTGGCGCCCAAGAGCTGATTCGTAATGCAAGTTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGAGCGCTGATCGAACAGACCGCTCTGTAACCGCAACAGCAAGGCGGATTC 720
Db 2151 TTGGGAGCGCTGATCGAACAGACCGCTCTGTAACCGCAACAGCAAGGCGGATTC 2210
Qy 721 CCCGCAACACATTCGAAACACCGGCTGATCCGCGCTTGAACCAAAATCAGCAGGGAAG 780
Db 2211 CCCGCAACACATTCGAAACACCGGCTGATCCGCGCTTGAACCAAAATCAGCAGGGAAG 2270
Qy 781 GAAAAACGCGGCAAGGCGCGCAACAGTTCATG 814
Db 2271 GAAAAACGCGGCAAGGCGCGCAACAGTTCATG 2304

RESULT 8
US-10-096-129-1
; Sequence 1, Application US/10096129
; Publication No. US20030207406A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: ROTH, STEPHEN
; APPLICANT: BUCZALA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO SACCHARIDE UNITS WITH A
; FILE REFERENCE: POLYGLYCOSYLTRANSFERASE
; FILE REFERENCE: 040853-01-5021-02
; CURRENT APPLICATION NUMBER: US/10/096,129
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/338,943
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 08/478,140
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: *Neisseria gonorrhoeae*
US-10-096-129-1

Query Match 87.7%; Score 726; DB 6; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,5e-232;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATACGCTTAGCTTCCGCGCAAGACGAGGCGGCAATTCGCCAT 60
Db 1491 ATGCAAAACCAAGTATACGCTTAGCTTCCGCGCAAGACGAGGCGGCAATTCGCCCA 1550

Qy 61 ACCTTGGCAGGACAGGATCCGTTTCAAGTTTTCAGACGACATGATCCGCTGTAAGG 120
Db 1551 ACCTTGGCAGGATCCGTTTCAAGTTTTCAGACGACATGATCCGCTGTAAGG 1610
Qy 121 CTGGAACAGGCAATGCGGAACTGATCCCGGCTTGTGCGGCAACCCCTATTTGAGCGGA 180
Db 1611 CTGGAACAGGCAATGCGGAACTGATCCCGGCTTGTGCGGCAACCCCTATTTGAGCGGA 1670
Qy 181 GTGAAAAAGCTGCTTTATGAGCCACGCGGTATGTTGAAAGCAGGCAATGAGCAGAGT 240
Db 1671 GTGAAAAAGCTGCTTTATGAGCCACGCGGTATGTTGAAAGCAGGCAATGAGCAGAGT 1730
Qy 241 CTGCGTATATACCGTATTTGAGGACAGCTTTTACTCGGCAAGGGAAGAAATTC 300
Db 1731 GTACGTTATGCGCGTATTTGAAAGATATGTTCTTACTCGGCAAGGCGGAGCAGCTTC 1790
Qy 301 CTTGCCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGCTCCG 360
Db 1791 CTTGCCGAAGATCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGCTCCG 1850
Qy 361 TTGGAACAGATGTTTATGACAGTCTGACCTGCGCTCCGCGGTGCGGATTAATCTCGCG 420
Db 1851 TTGGAACAGATGTTTATGACAGTCTGACCTGCGCTCCGCGGTGCGGATTAATCTCGCG 1910
Qy 421 CGCGCTTTCCGCTGTTGAAAGCGAACCTGCGGAGACGCGGCTATATGATTTCCGA 480
Db 1911 CGCGCTTTCCGCTTTTGAAGCGAACCTGCGGAGACGCGGCTATATTTCCGA 1970
Qy 481 AAGGCGATGCGGTTTTCCTGACAGGTTTTCGCGCTCCGCGCGGCAAGGCTGACCTCC 540
Db 1971 AAGGCGATGCGGTTTTCCTGACAGGTTTTCGCGCTCCGCGCGGCAAGGCTGACCTCC 2030
Qy 541 GTGCAATCTGATGATGTTTCAAGCGATTTTTCGACAGGGAAGATGCGGTTTGGCAGCTC 600
Db 2031 GTGCAATTTGATGATGTTTTCGCAACCTGACAGGGAAGATGCGGTTTGGCAGCTC 2090
Qy 601 AATCCCGCTTGGCGCCCAAGAGCTGATTCGTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCCGCTTGGCGCCCAAGAGCTGATTCGTAATGCAAGTTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGAGCGCTGATCGAACAGACCGCTCTGTAACCGCAACAGCAAGGCGGATTC 720
Db 2151 TTGGGAGCGCTGATCGAACAGACCGCTCTGTAACCGCAACAGCAAGGCGGATTC 2210
Qy 721 CCCGCAACACATTCGAAACACCGGCTGATCCGCGCTTGAACCAAAATCAGCAGGGAAG 780
Db 2211 CCCGCAACACATTCGAAACACCGGCTGATCCGCGCTTGAACCAAAATCAGCAGGGAAG 2270
Qy 781 GAAAAACGCGGCAAGGCGCGCAACAGTTCATG 814
Db 2271 GAAAAACGCGGCAAGGCGCGCAACAGTTCATG 2304

RESULT 9
US-10-654-528-1
; Sequence 1, Application US/10654528
; Publication No. US20040043464A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387

241 CTGCCGTATATCACCGTATTGTAGACGACGTTTACTCGGCGAAGGTGAGAAAAATTC 300

QY 301 CTTGCCGAGAGACGCTTGCTGCAAGACGCTTTGATCCCGGATACCGCTTTTATGTCGCCG 360
Db 1791 CTTGCCGAGAAATCTTGGCTGCAAGAACGCTTTTGACCCCGATTCGCTTTTGTGTCGCCG 1850
QY 361 TTGGAAGACGATGTTTATGACGCTGCTGACCTCGCCCTCGGCGGATTACTGCGGG 420
Db 1851 TTGGAAGACGATGTTTATGACGCTGCTGACCTCGCCCTCGGCGGATTACTGCGGG 1910
QY 421 CGGCGCTTTCCGCTGTTTGAAGACGACATCGGCGGACGCGCGCTTATCATTTTCCGA 480
Db 1911 CGGCGCTTTCCGCTGTTTGAAGACGACATCGGCGGACGCGCGCTTATCATTTTCCGA 1970
QY 481 AAGGACATGCGGTTTTCCTGCAAGGTTTCCGCGCTCGCGCGCAAGGCGTGCACCC 540
Db 1971 AAGGACATGCGGTTTTCCTGCAAGGTTTCCGCGCTCGCGCGCAAGGCGTGCACCC 2030
QY 541 GTGATCTGATGATGTTTCAAGCGGATTTTTCGACAGGAGAAAGATGCGGTTTGCAGCTC 600
Db 2031 GTGATCTGATGATGTTTCAAGCGGATTTTTCGACAGGAGAAAGATGCGGTTTGCAGCTC 2090
QY 601 AATCCCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 660
Db 2091 AATCCCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 2150
QY 661 TTGGGAGGCTGATGCAACGACGCGCTTCGAAACGCGAAAGGCGGCTTC 720
Db 2151 TTGGGAGGCTGATGCAACGACGCGCTTCGAAACGCGAAAGGCGGCTTC 2210
QY 721 CCGGCAACACATTTAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTTAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGGCAAGGCGCGCAAGCTTCAATTG 814
Db 2271 GAAAAACGCGGCAAGGCGCGCAAGCTTCAATTG 2304

RESULT 11
US-10-472-260-175
Sequence 175, Application US/10472260
Publication No. US20040265328A1
GENERAL INFORMATION:
APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS LIMITED
FILE REFERENCE: GMS/DJC/23480
CURRENT APPLICATION NUMBER: US/10/472,260
NUMBER OF SEQ ID NOS: 199
SOFTWARE: Patent version 3.1
SEQ ID NO 175
LENGTH: 840
TYPE: DNA
ORGANISM: *Neisseria meningitidis* (group B)
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(840)
OTHER INFORMATION: NMB1926
US-10-472-260-175

Query Match 51.4%; Score 426; DB 8; Length 840;
Best Local Similarity 75.2%; Pred. No. 6,6e-132;
Matches 531; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 1 ATGCAAAACACGTTATCAGTTAGCTTCCGCGGAGAACGAGGCGCACATTGCCGAT 60
Db 1 ATGCAAAACACGTTATCAGTTAGCTTCCGCGGAGAACGAGGCGCACATTGCCGAT 60
QY 61 ACCTTCGAGGACGCGCATCCCGTTGCTTTTCAGACGCTGATGCGGCTGAAAG 120
Db 61 ACCTTCGAGGACGCGCATCCCGTTGCTTTTCAGACGCTGATGCGGCTGAAAG 120

QY 121 CTGGAACAGGCAATGGCGAACTCGTCCCCTGCTGTGCGGCAACCCCTATTGAGCGGA 180
Db 121 CTGGAATCGGATGATGGCGAAATTGGTGCCTGGGATGGCAAAACAGCACCTGTGAGCGAA 180
QY 181 GTGAAAAAGCGCTGTTTATGAGCCAGCGCTATTGTGAAAGCGGCAATTGAGCAAGT 240
Db 181 GTGAAAAAGCGCTGTTTATGAGCCAGCGCTATTGTGAAAGCGGCAATTGAGCAAGT 240
QY 241 CTGCGTATATCACCGTATTTGAGGACGAGTTTACTCGGCGAAGGTGAGAAAAATTC 300
Db 241 CTGCGCTATGTTGCGTATTTGAGGATGATGCTCGCTTGGCAAGAGCAAGAAAGTTTC 300
QY 301 CTTGCCGAAGACGCTTGGCTGCAAGACGCTTGAACCGGATACCGCTTATGCTCGG 360
Db 301 CTTGCCGAAGATCTTGGCTGCAAGACGCTTGAATAGATTCGCTTTTATGCTTGT 360
QY 361 TTGGAAGACGATGTTTATGACGCTCCTGACCTCGCCCTCGGCGTGGCGGATTACTGCGGG 420
Db 361 TTGGAAGACGATGTTTATGACGCTCCTGACGCTGATTAAGTCTGAAATTAAGAGAT 420
QY 421 CGCGCTTTCCGCTGTGGAAGCGAACACTGGGGGACGCGGCGTATATCATTTCCCGA 480
Db 421 CGGCTATTTCTTGTGCTGAGAGCGAACATTTGGGGAGCGCTGGCTATATCATTTCCGCT 480
QY 481 AAGGACATGCGGTTTTCCTGCAAGGTTTCCGCGCTGCGCGCAAGGCGCTGCACCC 540
Db 481 AAGGACATGCGGTTTTCCTGCAAGGTTTCCGCGCTGCGCGCAAGGCGCTGCACCC 540
QY 541 GTGATCTGATGATGTTTCAAGCGGATTTTTCGACAGGAGAAAGATGCGGTTTGCAGCTC 600
Db 541 GTGATCTGATGATGTTTCAAGCGGATTTTTCGACAGGAGAAAGATGCGGTTTGCAGCTC 600
QY 601 AATCCCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 660
Db 601 AATCCCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 660
QY 661 TTGGGAGGCTGATGCAACGACGCGCTTCGAAACGCGCAAGCAGC 706
Db 661 TTGGGAGGCTGATGCAACGACGCGCTTCGAAACGCGCAAGCAGC 706

RESULT 12
US-10-795-159-570
Sequence 570, Application US/10795159
Publication No. US20050221439A1
GENERAL INFORMATION:
APPLICANT: BAKALITZ et al.
TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
FILE REFERENCE: 28335/38815A
CURRENT APPLICATION NUMBER: US/10/795,159
NUMBER OF SEQ ID NOS: 771
SOFTWARE: Patent version 3.2
SEQ ID NO 570
LENGTH: 14547
TYPE: DNA
ORGANISM: *H. influenzae*
FEATURE:
NAME/KEY: misc feature
LOCATION: (14474)..(14474)
OTHER INFORMATION: n = a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (14508)..(14508)
OTHER INFORMATION: n = a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (14518)..(14518)
OTHER INFORMATION: n = a, c, g, or t
US-10-795-159-570

Query Match 11.2%; Score 93; DB 9; Length 14547;
Best Local Similarity 51.9%; Pred. No. 1.3e-19;
Matches 225; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 180 AGGAGAAAAAGCTGCTTTATGAGCAAGCCGATTTGTGAGACAGGCAATTGACAGAG 239
DB 11201 AGGAGAAAAAGCTGCTTTATGAGCAATTCCTTATTATGAGAAATAGTGTGAGAAAA 11260

QY 240 TCGGCGGTATATACCGGTATTTAGAGACGCTTTACTCGGAGAGGTGAGAAAAAT 299
DB 11261 TTTGGAATATCTACAAATTTTGAAGATGATTAATTTCTGGAGAAATGCGAAGTGT 11320

QY 300 CCTTCCGAGACGCTTGTGCTGCAAGACGCTTTGACCCGATATACGCTTTATCTGCG 359
DB 11321 TTTGGACACAGATGAAATGTTAAAAACGCTTTGATTTTAATGATATTTTATTTG 11380

QY 360 CTTGGAAAGCATTTTATGACGT---CTTACCTCGCCCTCGGCGTGGCGGATTAATG 416
DB 11381 TTTAGAACTTTTTCAGCCAGTTAACTTGAGAAAACAAACTAAATTCACCTTTT 11440

QY 417 CGGCGCGGCTTCCGCTGTGGAAGCCAGACCTGGGGAGCGCGGCTATATCATTTTC 476
DB 11441 TTTCTAGAACTTTGATATTTTAAATTCAGCTCACTGGGGAGCGCTGTATATTTTC 11500

QY 477 CCGAAAGGAGATCGGCTTTTCTGGAAGATTTGCGCCGCTCGCCGAGGCGCTGCA 536
DB 11501 TCAAGGTGCGGCTTAATATGTAATTTAAATTTAAAGAAATATTCCTTGTGATAAAT 11560

QY 537 CCCCCTGATCTGATGATGATGTCAGCAATTTTTCAGACAGGAAGAAATGCGGTTGCCA 596
DB 11561 TCGAGTGTATGAATCTATTATTTTAAATTAATGATGATGATTAATATTTATCTATCA 11620

QY 597 GCTCATCCCGCTTGTGCGCCCAAGAGCTGCA 629
DB 11621 ACTTAATCCAGCAATTTGATTTCAAGAACTCCA 11653

RESULT 13
US-10-795-159-681
Sequence 681, Application US/10795159
Publication No. US20050221439A1
GENERAL INFORMATION:
APPLICANT: BAKALETZ et al.
TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
FILE REFERENCE: 28335/38815A
CURRENT FILING DATE: 2004-03-05
PRIOR APPLICATION NUMBER: US 60/453,134
PRIOR FILING DATE: 2003-03-06
NUMBER OF SEQ ID NOS: 771
SOFTWARE: Patentin version 3.2
SEQ ID NO 681
LENGTH: 106645
TYPE: DNA
ORGANISM: H. influenzae
US-10-795-159-681

Query Match 11.2%; Score 93; DB 9; Length 106645;
Best Local Similarity 51.9%; Pred. No. 2.7e-19;
Matches 225; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 180 AGGAGAAAAAGCTGCTTTATGAGCAAGCCGATTTGTGAGACAGGCAATTGACAGAG 239
DB 98869 AGGAGAAAAAGCTGCTTTATGAGCAATTCCTTATTATGAGAAATAGTGTGAGAAAA 98928

QY 240 TCGGCGGTATATACCGGTATTTAGAGACGCTTTACTCGGAGAGGTGAGAAAAAT 299
DB 98929 TTTGGAATATCTACAAATTTTGAAGATGATTAATTTCTGGAGAAATGCGAAGTGT 98988

QY 300 CCTTCCGAGACGCTTGTGCTGCAAGACGCTTTGACCCGATATACGCTTTATCTGCG 359
DB 98989 TTTGGACACAGATGAAATGTTAAAAACGCTTTGATTTTAATGATATTTTATTTTCG 99048

QY 360 CTTGGAAGCATGTTTATGACGT---CTTGAACCTCGCCCTCGGCGTGGCGGATTAATG 416
DB 99049 TTTAGAACTTTTATGAGCCAGTTAACTTGAGAAACAAATTCACCTTTT 99108

QY 417 CGGCGCGGCTTTCGCTGTTGAAAGCCAGACATCGGGGAGCGCGGCTATATCATTTTC 476
DB 99109 TTTCTAGAACTTTGATATTTTAAATTCAGCTCACTGGGGAGCGGCTGTATATTTTC 99168

QY 477 CCGAAAGGATGCGGCTTTTCTGGAAGATTTTCCGCTGCGCCCTCGGAGGCTGCA 536
DB 99169 TCAAGGTGCGGCTTAATATGTAATTTAAAGATATTTCTTCTGATGAAATTTGT 99228

QY 537 CCCCCTGATCTGATGATGTCAGCAATTTTTCAGACAGGAAGAAATCGGTTTGCA 596
DB 99229 TCGAGTGTATGAATCTATTATTTTAAATTAATTAATTAATTAATTAATTAATTA 99288

QY 597 GCTCAATCCCGCTTGTGCGCCCAAGAGCTGCA 629
DB 99289 ACTTAATCCAGCAATTTGATTTCAAGAACTCCA 99321

RESULT 14
US-10-329-670-1/c
Sequence 1, Application US/10329670
Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischnann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra
FILE REFERENCE: PB186P1
CURRENT FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US/10/329,670
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc.feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)

```
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152530)..(152530)
/
Query Match      10.5%; Score 86.6; DB 7; Length 1830121;
Best Local Similarity 51.0%; Pred. No. 1,1e-16;
Matches 231; Conservative 0; Mismatches 219; Indels 3; Gaps 1;
/
QY 180 AGTGAAGAAAAGCTGCTTTATGAGCCAGCCGCTATGTTGAGACAGGCAATTGACGAGG 239
/      |||||
DB 570747 AGGAGAAAAGGCTTTATGAGCCAGCTTTATGAGACAGGCAATTGACGAGG 570688
/
QY 240 TCTGCGGTATATGACCGGTATTTGAGGAGCGTTTATCTGCGGAGAGTGAAGAAAATT 299
/      |||||
DB 570687 TTTGATATATCTCAAAATTTTGAAGATGATATTTCTTGCGAGAAATCGAGATTT 570628
/
QY 300 CCTTGCAGAAAGCCTTGCTGCAAGAACGCTTTGACCCGATACCGCTTTATCTGCTCG 359
/      |||||
DB 570627 TTTGAACCAAAATGAATGTTAAACAGCTTTGATTTAATGATATTTTATATGCG 570568
/
QY 360 CTTGGAAGACATGTTATGACGT---CTGACCTGCGCTCCGCGGTGGCGGATTACTG 416
/      |||||
DB 570567 TTTAGAAACTTTTTCAGCCAGCTTAACTTGAGAAACAAACTTAAATTCACCTTTTAA 570508
/
QY 417 CGGCGCGCGCTTCCGCTGTTGGAAGCAACACTGGGGAGCGCGCGCTATATCACTTC 476
/      |||||
DB 570507 TTTTAGAAGACTTTGATATTTTAAATCACTGAGGGAGCGCGGTTATATTTTC 570448
/
QY 477 CCGAAAAGCGATCGGTTTTTCTGACAGGTTTGCGCGCCGCGCCGAGAGGCTGCA 536
/      |||||
DB 570447 TCAAGGTGGGCTTAAATATATATGATTTTAAAGATATTTCTTCTGATGAATGTG 570388
/
QY 537 CCCCCTGATCTGATGATGTTCCAGCATTTTTCGACAGGAAAGAAATCGGTTGCCA 596
/      |||||
DB 570387 TGCAGTGTGATGAACTTATTTTATTAATAGTGTGATTAATTAATTTGCTATCA 570328
/
QY 597 GCTCATCCCGCTTGCGGCGCCGAGAGCTGCA 629
/      |||||
DB 570327 ACTGAATCCAGCAATTTGATTTCAAGACTCCA 570295
/
RESULT 15
US-10-158-865-1/c
/ Sequence 1, Application US/10158865
/ Publication No. US20040203093A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleischmann et al.
/ TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
/ TITLE OF INVENTION: Thereof, and Uses Thereof
/ FILE REFERENCE: PB186P2C1D1
/ CURRENT APPLICATION NUMBER: US/10/158,865
/ CURRENT FILING DATE: 2002-06-03
/ PRIOR APPLICATION NUMBER: US 09/557,884
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 08/476,102
/ PRIOR FILING DATE: 1995-06-07
/ PRIOR APPLICATION NUMBER: US 08/426,787
/ PRIOR FILING DATE: 1995-04-21
/ NUMBER OF SEQ ID NOS: 1
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1830121
/ TYPE: DNA
/ ORGANISM: Haemophilus influenzae
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (4747)..(4747)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (9921)..(9921)
/ OTHER INFORMATION: n equals a,t,c, or g
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (10150)..(10150)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (29298)..(29298)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (36543)..(36543)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (36551)..(36551)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (36636)..(36636)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (40810)..(40810)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (44416)..(44416)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (44905)..(44905)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (44975)..(44975)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (45593)..(45593)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (45732)..(45732)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (47036)..(47036)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51334)..(51334)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51602)..(51602)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (51786)..(51786)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (55369)..(55369)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (65309)..(65309)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
```

```
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

Query Match 10.5%; Score 86.6; DB 8; Length 1830121;
Best Local Similarity 51.0%; Pred. No. 1.1e-16;
Matches 231; Conservative 0; Mismatches 219; Indels 3; Gaps 1;

```
QY 180 AGTGAAAAAGCGCTGTTATGAGCCAGCGATTTGGAAGCAGGATTTGAGCAAG 239
DB 570747 AGGAGAAAAAGCGTGTATTAATGACCTTATTAATGAGTAAGTGTAAGAAA 570688
QY 240 TCTGCCGTATACACCGATTTGAGAGAGCGTTTACCGCGAAGTGAGGAAAT 299
DB 570687 TTTGAAATATCTCAAAATTTTGAAGATGATTAATCTTGCGAGAAATCGAAAGTAT 570628
QY 300 CTTGCGAAGACGCTTGCGCTGCAAGAACGCTTTGACCCGATACCGCTTTATGTCG 359
DB 570627 TTTGAACCAAAATGATGTTAAACACGTTTGATTTATGATTAATTTATTTG 570568
QY 360 CTTGAAAGCATGTTATGACGT--CCTGACCTGCGCCCGCGGCGGATTAACG 416
DB 570567 TTTAGAACTTTTACGCGCATTAACCTTGAGAAACAACTAAATTCACCTTTAA 570508
QY 417 CGGGCGCGCTTTCGCTGTGGAAGCAACCTGCGGAGCGCGGCTATATCATTTTC 476
DB 570507 TTTCTAGGAACCTTGATATTTTAAATCACTGCGGCGGACGATTAATTTTTC 570448
QY 477 CGAAAGCATGCGGTTTTCCTGACAGGTTTGCCTGCGCCCGCGAAGGCTGCA 536
DB 570447 TCAAGTGGCGGTAATATGTAATTAATTAAGAAATATCTTCTGATGAATTC 570388
QY 537 CCCGCTGATCTGATGATGTTGAGCGATTTTTCACAGGAGGAAGCCGTTGCCA 596
DB 570387 TGCAGTTGATGAACCTTAATTTTAATTAATTAATTAATTAATTAATTAATTA 570328
QY 597 GCTCAATCCGCGCTTGCGCCCAAGAGCTGCA 629
DB 570327 ACTGAATCAGCAATTTGATTAACAAGACTCA 570295
```

Search completed: April 7, 2006, 18:56:47
Job time : 1367.05 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 1 7, 2006, 18:05:06 ; Search time 966.604 Seconds
(without alignments)
3426.906 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacacacgtatcag.....tcattgtgccttccaataa 828

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Maximum Match 0%
Listing first 45 summaries

Database :

Published Applications NA New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	726	87.7	5859	14 US-11-102-497-1	Sequence 1, Appli
2	726	87.7	5859	14 US-11-102-497-7	Sequence 7, Appli
3	721.2	87.1	837	8 US-10-467-657-1309	Sequence 1309, Ap
4	471.8	57.0	840	8 US-10-467-657-1325	Sequence 1325, Ap
5	57	6.9	2449	14 US-11-194-246-26	Sequence 26, Appl
6	38	4.6	1760	14 US-11-024-959-37	Sequence 37, Appl
7	34.6	4.2	453	8 US-10-997-437A-9	Sequence 9, Appli
8	34.6	4.2	1218	8 US-10-750-185-5685	Sequence 5685, A
9	34.6	4.2	1218	8 US-10-750-623-5685	Sequence 5685, A
10	33.8	4.1	11864	9 US-10-330-773-826	Sequence 826, App
11	33.6	4.1	535	9 US-09-925-065A-52784	Sequence 52784, A
12	33.6	4.1	535	9 US-10-301-480-154022	Sequence 154022,
13	33.6	4.1	535	10 US-10-301-480-767431	Sequence 767431,
14	33.6	4.1	7364	14 US-11-096-281-10	Sequence 10, Appl
15	33.4	4.0	626	6 US-09-925-065A-525973	Sequence 525973,
16	33.4	4.0	626	6 US-09-925-065A-525974	Sequence 525974,
17	33.4	4.0	626	6 US-09-925-065A-717818	Sequence 717818,
18	33	4.0	411	8 US-10-467-657-6075	Sequence 6075, Ap

ALIGNMENTS

```
C 19 33 4.0 555 8 US-10-467-657-6055 Sequence 6055, Ap
C 20 33 4.0 555 8 US-10-467-657-7535 Sequence 7535, Ap
C 21 33 4.0 938 11 US-11-096-568A-18718 Sequence 18718, A
C 22 32.8 4.0 1866 8 US-10-467-657-1345 Sequence 1345, Ap
C 23 32.8 3.9 1694969 7 US-10-506-454-1690 Sequence 1690, Ap
C 24 32.4 3.9 1230 8 US-10-467-657-4857 Sequence 4857, Ap
C 25 32.2 3.9 6021 14 US-11-136-527-274 Sequence 274, App
C 26 32.2 3.9 1878 8 US-10-467-657-6425 Sequence 6425, Ap
C 27 32.2 3.9 1878 8 US-10-467-657-7617 Sequence 7617, Ap
C 28 32 3.9 834 8 US-10-750-185-37550 Sequence 37550, A
C 29 32 3.9 834 8 US-10-750-623-37550 Sequence 37550, A
C 30 32 3.9 1890 8 US-10-750-185-61161 Sequence 61161, A
C 31 32 3.9 1890 8 US-10-750-623-61161 Sequence 61161, A
C 32 32 3.9 2205 8 US-10-467-657-3729 Sequence 3729, Ap
C 33 32 3.9 2205 8 US-10-467-657-6989 Sequence 6989, Ap
C 34 31.8 3.8 808 11 US-11-096-568A-12049 Sequence 12049, A
C 35 31.8 3.8 1263 8 US-10-858-730-149 Sequence 149, App
C 36 31.6 3.8 1999 11 US-11-096-568A-14677 Sequence 14677, A
C 37 31.6 3.8 4734 14 US-11-136-527-2518 Sequence 2518, Ap
C 38 31.4 3.8 861 14 US-11-137-465-15 Sequence 15, Appl
C 39 31.4 3.8 165857 14 US-11-121-086-34 Sequence 34, Appl
C 40 31.2 3.8 1026 8 US-10-750-185-57052 Sequence 57052, A
C 41 31.2 3.8 1026 8 US-10-750-623-57052 Sequence 57052, A
C 42 31.2 3.8 7011 14 US-11-136-527-3740 Sequence 3740, Ap
C 43 31.2 3.8 9695 14 US-11-096-281-12 Sequence 12, Appl
C 44 31.2 3.8 88421 14 US-11-205-109-1 Sequence 1, Appli
C 45 31 3.7 1348 9 US-10-838-616-33 Sequence 33, Appl

RESULT 1
US-11-102-497-1
; Sequence 1, Application US/1102497
; Publication No. US20050271690A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C
; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(381)
; OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (445)..(1491)
; OTHER INFORMATION: 19CA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2342)..(3262)
; OTHER INFORMATION: 19rc
```

```
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (3322)..(4335)
/ OTHER INFORMATION: 1gtd
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (4354)..(5196)
/ OTHER INFORMATION: 1gce
US-11-102-497-1
```

```
Query Match      87.7%; Score 726; DB 14; Length 5859;
Best Local Similarity 93.2%; Pred. No. 3e-215;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
```

```
QY 1 ATGCAAAACACGTTATCAAGCTTAGCTTCGCGCGAGAAACGCGGCGGCAATTGCCGAT 60
DB 1491 ATGCAAAACACGTTATCAAGCTTAGCTTCGCGCGAGAAACGCGGCGGCAATTGCCGCA 1550
QY 61 ACCTTGCGGAGGCAAGCGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGTGAAAG 120
DB 1551 ACCTTGCGGAGGCGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGTGAAAG 1610
QY 121 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTTCGCGGCAATCCCTATTTCAGCGGA 180
DB 1611 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTTCGCGGCAATCCCTATTTCAGCGGA 1670
QY 181 GTGGAAAAAGCGCTTTATGAGCGCAAGCGGTAATGTTGGAAGAGGCAATTGACGAAAGT 240
DB 1671 GTGGAAAAAGCGCTTTATGAGCGCAAGCGGTAATGTTGGAAGAGGCAATTGACGAAAGC 1730
QY 241 CTGCGGTATATCAACGCTATTTCAGAGCAAGTTTTCAGGCGAAGTGAAGAAATTC 300
DB 1731 GTAACCGTAATACGCGCTATTTCAGAGCAAGTTTTCAGGCGAAGTGAAGAAATTC 1790
QY 301 CTTCGGAAGACGCTTTCGCTGCAAGACGCTTTCAGACCGGTAATGCGCTTTCAGCGGC 360
DB 1791 CTTCGGAAGACGCTTTCGCTGCAAGACGCTTTCAGACCGGTAATGCGCTTTCAGCGGC 1850
QY 361 TTGGAAACGATGTTTATGACAGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 420
DB 1851 TTGGAAACGATGTTTATGACAGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 1910
QY 421 CGCGCTTTCGCTGTTGGAAGAGCACTGCGGCGGCAAGCGGCTATATATTCGCGA 480
DB 1911 CGCGCTTTCGCTGTTGGAAGAGCACTGCGGCGGCAAGCGGCTATATATTCGCGA 1970
QY 481 AAAGCGATGCGGTTTCTTGACAGGTTTTCGACGCGCTTCGCGGCGGCAAGCGGCTGACCC 540
DB 1971 AAAGCGATGCGGTTTCTTGACAGGTTTTCGACGCGCTTCGCGGCGGCAAGCGGCTGACCC 2030
QY 541 GTTCGATCTGATGATGTTTCAGCGATTTTTCGACAGGGAAGAAATGCGGTTTTCGAGCTC 600
DB 2031 GTTCGATCTGATGATGTTTCAGCGATTTTTCGACAGGGAAGAAATGCGGTTTTCGAGCTC 2090
QY 601 AATCCGCTTTCGCGGCGGCAAGGCTGATTAATGCGAAGTTTCGAGCAAGCAAGCGGA 660
DB 2091 AATCCGCTTTCGCGGCGGCAAGGCTGATTAATGCGAAGTTTCGAGCAAGCAAGCGGA 2150
QY 661 TTGGGAGGCTGATGCAACGAGCGCTTCGAAACCGCAACAGCAAGCGGCAATTC 720
DB 2151 TTGGGAGGCTGATGCAACGAGCGCTTCGAAACCGCAACAGCAAGCGGCAATTC 2210
QY 721 CCGCGCAACATTCGAAACAGCGGCTGATCCGCGCTTTCGCAAAATGCAAGGAAAGG 780
DB 2211 CCGCGCAACATTCGAAACAGCGGCTGATCCGCGCTTTCGCAAAATGCAAGGAAAGG 2270
QY 781 GAAAAACGCGGCAAGGCGGCAAGGCTGATTC 814
DB 2271 GAAAAACGCGGCAAGGCGGCAAGGCTGATTC 2304
```

RESULT 2
US-11-102-497-7

```
/ Sequence 7, Application US/11102497
/ Publication No. US20050271690A1
/ GENERAL INFORMATION:
/ APPLICANT: Gotschlich, Emil C.
/ TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
/ FILE REFERENCE: 040853-01-5029-03/102,497
/ CURRENT FILING DATE: 2005-04-08
/ PRIOR APPLICATION NUMBER: US/10/654,528
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 10/007,267
/ PRIOR FILING DATE: 2001-12-03
/ PRIOR APPLICATION NUMBER: US 09/333,412
/ PRIOR FILING DATE: 1999-06-15
/ PRIOR APPLICATION NUMBER: US 08/878,360
/ PRIOR FILING DATE: 1997-06-18
/ PRIOR APPLICATION NUMBER: US 08/683,426
/ PRIOR FILING DATE: 1996-07-18
/ PRIOR APPLICATION NUMBER: US 08/312,387
/ PRIOR FILING DATE: 1994-09-24
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 7
/ LENGTH: 5859
/ TYPE: DNA
/ ORGANISM: Neisseria gonorrhoeae
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1491)..(2330)
/ OTHER INFORMATION: 1gcb
US-11-102-497-7
```

```
Query Match      87.7%; Score 726; DB 14; Length 5859;
Best Local Similarity 93.2%; Pred. No. 3e-215;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
```

```
QY 1 ATGCAAAACACGTTATCAAGCTTAGCTTCGCGCGAGAAACGCGGCGGCAATTGCCGAT 60
DB 1491 ATGCAAAACACGTTATCAAGCTTAGCTTCGCGCGAGAAACGCGGCGGCAATTGCCGCA 1550
QY 61 ACCTTGCGGAGGCAAGCGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGTGAAAG 120
DB 1551 ACCTTGCGGAGGCGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGTGAAAG 1610
QY 121 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTTCGCGGCAATCCCTATTTCAGCGGA 180
DB 1611 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTTCGCGGCAATCCCTATTTCAGCGGA 1670
QY 181 GTGGAAAAAGCGCTTTATGAGCGCAAGCGGTAATGTTGGAAGAGGCAATTGACGAAAGT 240
DB 1671 GTGGAAAAAGCGCTTTATGAGCGCAAGCGGTAATGTTGGAAGAGGCAATTGACGAAAGC 1730
QY 241 CTGCGGTATATCAACGCTATTTCAGAGCAAGTTTTCAGGCGAAGTGAAGAAATTC 300
DB 1731 CTGCGGTATATCAACGCTATTTCAGAGCAAGTTTTCAGGCGAAGTGAAGAAATTC 1790
QY 301 CTTCGGAAGACGCTTTCGCTGCAAGACGCTTTCAGACCGGTAATGCGCTTTCAGCGGC 360
DB 1791 CTTCGGAAGACGCTTTCGCTGCAAGACGCTTTCAGACCGGTAATGCGCTTTCAGCGGC 1850
QY 361 TTGGAAACGATGTTTATGACAGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 420
DB 1851 TTGGAAACGATGTTTATGACAGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 1910
QY 421 CGCGCTTTCGCTGTTGGAAGAGCACTGCGGCGGCAAGCGGCTATATATTCGCGA 480
DB 1911 CGCGCTTTCGCTGTTGGAAGAGCACTGCGGCGGCAAGCGGCTATATATTCGCGA 1970
QY 481 AAAGCGATGCGGTTTCTTGACAGGTTTTCGACGCGCTTCGCGGCGGCAAGCGGCTGACCC 540
DB 1971 AAAGCGATGCGGTTTCTTGACAGGTTTTCGACGCGCTTCGCGGCGGCAAGCGGCTGACCC 2030
```

```
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCAGAGGAAAGTCCGTTTCCAGCTC 600
    |||||
Db 2031 GTGATCTGATGATGTTTTCAGCGATTTTTCAGAGGAAAGTCCGTTTCCAGCTC 2090
QY 601 AATCCCGCTTTGCGCCCAAGAGCTGATTAATGCGAATTTTCAAGCAAAACAGCGCA 660
    |||||
Db 2091 AATCCCGCTTTGCGCCCAAGAGCTGATTAATGCGAATTTTCAAGCAAAACAGCGCA 2150
QY 661 TTGGGAGAGCTGATTCAGACGACGCGCTTCGTAACCGCAAAACAGCAAGCGCGATTC 720
    |||||
Db 2151 TTGGGAGAGCTGATTCAGACGACGCGCTTCGTAACCGCAAAACAGCAAGCGCGATTC 2210
QY 721 CCGGCCAACACATTTCAAAACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 780
    |||||
Db 2211 CCGGCCAACACATTTCAAAACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGGCAAGGCGCGAACAGTTGATG 814
    |||||
Db 2271 GAAAAACGCGGCAAGGCGCGAACAGTTGATG 2304
```

```
RESULT 3
US-10-467-657-1309
; Sequence 1309, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMin99, version 1.04
; SEQ ID NO 1309
; LENGTH: 817
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1309
```

```
Query Match 87.1%; Score 721.2; DB 8; Length 817;
Best Local Similarity 92.9%; Pred. No. 5.2e-214;
Matches 756; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
QY 1 ATGCAAAACCAAGTTATCAGCTTACGTTCCGCGCAAGACGCGGCGCATTTGCCGAT 60
    |||||
Db 1 ATGCAAAACCAAGTTATCAGCTTACGTTCCGCGCAAGACGCGGCGCATTTGCCGAT 60
QY 61 ACCCTGGGCAAGGACGCGCATCCGCTTTCAGTTTTCGAGCAGCATGATCCGCTGAAG 120
    |||||
Db 61 ACCCTGGGCAAGGACGCGCATCCGCTTTCAGTTTTCGAGCAGCATGATCCGCTGAAG 120
QY 121 CTGGAACAGGCAATGGCGAATCGTCCCGGCTTTCGCGGCAACCCCTATTGAGCGGA 180
    |||||
Db 121 CTGGAACAGGCAATGGCGAATCGTCCCGGCTTTCGCGGCAACCCCTATTGAGCGGA 180
QY 181 GTGGAAAAAGCCGCTTATGAGCAGCGCGTATTTGGAAGCAGGATGGAAGAGT 240
    |||||
Db 181 GTGGAAAAAGCCGCTTATGAGCAGCGCGTATTTGGAAGCAGGATGGAAGAGT 240
QY 241 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTACTCGCGCAAGAGTGAAGAAATTC 300
    |||||
Db 241 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTACTCGCGCAAGAGTGAAGAGTTC 300
QY 301 CTTCGGAAGAGCGCTTGGCTGCAAGAACGCTTTCAGCCGATACCGCTTTATCGTCCG 360
    |||||
Db 301 CTTCGGAAGAGCGCTTGGCTGCAAGAACGCTTTCAGCCGATACCGCTTTATCGTCCG 360
```

```
QY 361 TTGAAACGATGTTTATGACGCTCTGACCTCGCTCCGCGTGGCGGATTAAGTCGGG 420
    |||||
Db 361 TTGAAACGATGTTTATGACGCTCTGACCTCGCTCCGCGTGGCGGATTAAGTCGGG 420
QY 421 CCGGCTTTCCGCTGTGGAAAAAGCACTGCGGAGACGCGCGCTATATCATTTCCGA 480
    |||||
Db 421 CCGGCTTTCCGCTGTGGAAAAAGCACTGCGGAGACGCGCGCTATATTTCCGA 480
QY 481 AAGGCAATGCGGTTTCTCTGACAGGTTTTCGCGCTCCGCGCGCGGCTGACCC 540
    |||||
Db 481 AAGGCAATGCGGTTTCTCTGACAGGTTTTCGCGCTCCGCGCGCGGCTGACCC 540
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCAGAGGAAAGATGCGGTTTCCAGCTC 600
    |||||
Db 541 GTGATCTGATGATGTTTTCAGCGATTTTTCAGAGGAAAGATGCGGTTTCCAGCTC 600
QY 601 AATCCCGCTTTGCGCCCAAGAGCTGATTAATGCGAATTTTCAGACCAAAACAGCGCA 660
    |||||
Db 601 AATCCCGCTTTGCGCCCAAGAGCTGATTAATGCGAATTTTCAGACCAAAACAGCGCA 660
QY 661 TTGGGAGAGCTGATGGAACAGACGCGCTCTGAAACCGCAACAGCAAGGCGCATTC 720
    |||||
Db 661 TTGGGAGAGCTGATGGAACAGACGCGCTCTGAAACCGCAACAGCAATGCGCGCATTC 720
QY 721 CCGGCCAACACATTTCAAAACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 780
    |||||
Db 721 CCGGCCAACACATTTCAAAACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 780
QY 781 GAAAAACGCGGCAAGGCGCGAACAGTTGATG 814
    |||||
Db 781 GAAAAACGCGGCAAGGCGCGAACAGTTGATG 814
```

```
RESULT 4
US-10-467-657-1325
; Sequence 1325, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMin99, version 1.04
; SEQ ID NO 1325
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1325
```

```
Query Match 57.0%; Score 471.8; DB 8; Length 840;
Best Local Similarity 80.7%; Pred. No. 2.2e-136;
Matches 551; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
```

```
QY 1 ATGCAAAACCAAGTTATCAGCTTACGTTCCGCGCAAGACGAGGCGCGCATTTGCCGAT 60
    |||||
Db 1 ATGCAAAACCAAGTTATCAGCTTACGTTCCGCGCAAGACGAGGCGCGCATTTGCCGCA 60
QY 61 ACCCTGGGCAAGGACGCGCATCCGCTTTCAGTTTTCGAGCAGCATGATCCGCTGAAG 120
    |||||
Db 61 ACCCTGGGCAAGGACGCGCATCCGCTTTCAGTTTTCGAGCAGCATGATCCGCTGAAG 120
QY 121 CTGGAACAGGCAATGGCGAATCGTCCCGGCTTTCGCGGCAACCCCTATTGAGCGGA 180
    |||||
Db 121 CTGGAACAGGCAATGGCGAATCGTCCCGGCTTTCGCGGCAACCCCTATTGAGCGGA 180
```

QY	181	GTGAAAAAAGCTCTCTTATGAGCCACGCCCTATTTGTGAAGACAGGCATTGGAAGAAAGT	240
Db	181	GTGAAAAAAGCTCTCTTATGAGCCACGCCCTATTTGTGAAGACAGGCATTGGAAGAAAGT	240
QY	241	CTGGCGTATATACCGTATTTGAGGACGACGTTTAACTCGCGCGAAGGAGAGAAAAATTC	300
Db	241	CTGGCGTATATACCGTATTTGAGGACGACGTTTAACTCGCGCGAAGGAGAGAAAAATTC	300
QY	301	CTTGCCGAAAGCGCTTGCTGCAAGAACGCTTTGACCCGATACCGCCTTATGTCGCC	360
Db	301	CTTGCCGAAAGCGCTTGCTGCAAGAACGCTTTGACCCGATACCGCCTTATGTCGCC	360
QY	361	TTGAAAAAGATGTTTATGACGTCCTGACCTCGCCTCCGCGTGGCGAGATTACCTGCGGG	420
Db	361	TTGAAAAAGATGTTTATGACGTCCTGACCTCGCCTCCGCGTGGCGAGATTACCTGCGGG	420
QY	421	CGGCGCTTTCCGCTGTTGGAAGGACAACCTGGGGGAGCGGGGCTATATATCTTCCGA	480
Db	421	CGGCGCTTTCCGCTGTTGGAAGGACAACCTGGGGGAGCGGGGCTATATATCTTCCGA	480
QY	481	AAAGCGATGCGGTTTTTCTGGAACAAGTTTCCGCGCTGCGCCGGAAGGCTGCACCC	540
Db	481	AAAGCGATGCGGTTTTTCTGGAACAAGTTTCCGCGCTGCGCCGGAAGGCTGCACCC	540
QY	541	GTCGATCTGATGATGTCAGCGATTTTTTTCACAGGGAAGAAATGCCGGTTGCGAGCTC	600
Db	541	GTCGATCTGATGATGTCAGCGATTTTTTTCACAGGGAAGAAATGCCGGTTGCGAGCTC	600
QY	601	AATCCCGGCTTGTGGGCCAAGAGCTGCATATGCAAAATTTCAGACGAAACGACGCA	660
Db	601	AATCCCGGCTTGTGGGCCAAGAGCTGCATATGCAAAATTTCAGACGAAACGACGCA	660
QY	661	TTGGGCAAGCTGATCGAAACGCA	683
Db	661	TTGGGCAAGCTGATCGAAACGCA	683

```

RESULT 5
US-11-194-246-26
; Sequence 26, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Molt, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Stefan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME
; TITLE OF INVENTION: USE
; FILE REFERENCE: 00592.US1 (MER 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 2449
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Sequence of the ribB coding sequence and flanking regions.
; US-11-194-246-26

```

D _b	1632	ATTGGTAAACAGAAATATCTCTTTCGAATTTTTTGATGCTATTACGCCCGCAGATTATTGA	1691
Q _Y	123	GGAAACAGAGC---AATGGCGGAACCTGCTCCCGGCTGTGTTCGGCGGACACCCCTATTTTGAGCGG	179
D _b	1692	AGAAACCCGCTAAAAAATTATATATTACATTAGATCGCTCTCTTAAAGCCAAAGTTGTGGGA	1751
Q _Y	180	AGTGGAAAAAGCTGCTTTATGAGCCACGCCCGTATGTGAAAGCAGGCATTGACGAAAG	239
D _b	1752	TGGGGAATAGGTGTGTCATTAAACCATATTGTTTATGGGATTTAGCATTTAGAAAAATPA	1811
Q _Y	240	TCTGCCGTATATCACCCGTATTTTGAAGACGAGTTTACTCGGCCGAGGTGAGAAAAATT	299
D _b	1812	TTTAACTATATCAATATCTTTGAAAGATGATATTCATTTGGGGAAAAATGCCAAGAATT	1871
Q _Y	300	CCCTTG 304	
D _b	1872	ATTGAG 1876	

```

RESULT 6
US-11-024-959-37
Sequence 37, Application US/11024959
Publication No. US2006001051eA1
GENERAL INFORMATION:
APPLICANT: FORSTER, RICHARD L.
APPLICANT: CONNETT, MARIE B.
APPLICANT: EMERSON, SARAH JANE
APPLICANT: GRIGOR, MURRAY ROBERT
APPLICANT: HIGGINS, COLLEEN M.
APPLICANT: LUND, STEVEN TROY
APPLICANT: MAEUSIN, ANDREAS
APPLICANT: KORZECZKI, BOB
TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
FILE REFERENCE: 044463-0360
CURRENT APPLICATION NUMBER: US/11/024,959
CURRENT FILING DATE: 2004-12-30
PRIOR APPLICATION NUMBER: 60/553,036
PRIOR FILING DATE: 2003-12-30
NUMBER OF SEQ ID NOS: 742
SOFTWARE: PatentIn version 3.3
SEQ ID NO 37
LENGTH: 1760
TYPE: DNA
ORGANISM: Eucalyptus sp.
US-11-024-959-37

```

Query Match	4.6%;	Score 38;	DB 14;	Length 1760;
Best Local Similarity	50.5%;	Pred. No. 0.28;		
Matches 92;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;

QY	590	TTTGCCAGCTCAATCCCGCCTTGTCGCGCCCAAGAGCTGCATTATGCAAGTTTACGACC	649
Db	123	TCACAGCGCCGAGCGATGGCGACTACTACGGGCGAGGCGACCCGATGAGCCCA	182
QY	650	AAAACAGCGCATTTGGGCGGCTGATGGAACGACGCGCTTCGTGAACCGCAACAGCAAA	709
Db	183	GGATCCGGATGGGCGACAACTTCATGCTCCACTACCTCCACCGCGCATGGAGATCA	242
QY	710	GGGCGCATTTCCCGCGCCCAACACATTCAACACGCGCTGATCCGGCGCTTGACCAAAATCA	769
Db	243	GCGGCGCCCTTCCCGCGCGCACACGACGACATCGCGCCTTCACACTCGAGAGACTACGTCA	302
QY	770	GC 771	
Db	303	CC 304	

RESULT 7
US-10-97-437A-9/c
; Sequence 9, Application US/1097437A
; Publication No. US20050250184A1
; GENERAL INFORMATION:

```
APPLICANT: Sheppard, Paul O.
APPLICANT: Fox, Brian A.
TITLE OF INVENTION: Peptide Hormones Zalpha48 and Zsig97
FILE REFERENCE: 03-16
CURRENT APPLICATION NUMBER: US/10/997,437A
CURRENT FILING DATE: 2004-11-24
PRIOR APPLICATION NUMBER: 60/525,597
PRIOR FILING DATE: 2003-11-26
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 453
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: degenerate sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(453)
OTHER INFORMATION: N = A,T,C, or G; W = A or T; S = G or C; Y = T or
OTHER INFORMATION: C; M = A or C; R = G or A; H = A or C or T
US-10-997-437A-9
```

```
Query Match          4.2%; Score 34.6; DB 8; Length 453;
Best Local Similarity 25.2%; Pred. No.2.2;
Matches 70; Conservative 47; Mismatches 161; Indels 0; Gaps 0;
```

```
QY 284 AAGGTGAGAAAATTCCTTCCGGAAGACGCTTGCTGCAAGAACGCTTGACCCGATA 343
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 353 ARCKRTGTAARGTGTTNSWRCAVTTNGNSWRAAFANANNGNCCNGTNAARTGYTTN 294
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 344 CGGCTTTATCGTCCGCTTGGAACGATGTTTNGACGCTCGACTCGCCCTCCGGCG 403
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 293 ARPAATYTRCTCTTCATNCNARNTCCNKGNAACDTTTCNACNCTYTGTCNGNSWN 234
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 404 TGGCGATTACTCGCGGCGCGCTTTCGCTGTGGAACGAACTGGGGAGCGCGG 463
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 233 GGNCCNARNCCNGCYTTCGNCNCKNCCNARNGCCARCTCNCKNCCNANATYGNARN 174
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 464 GCTATATCATTTCCGGAAGACGATGGGTTTTTCTCGACAGGTTTGGCGCCCTGCGC 523
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 173 CCYTTTGTGTCNCGNSRTRGTGTTTCKNARYTCYGNACNARYTCNACNARNCKN 114
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 524 CCGAAGGCTGCAACCCCTGATCTGATGATGTTTCCAGC 561
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 113 ARNARNGCTYGNCCRTCCNCGGTCNCKNGGTYTCNGC 76
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 8

```
US-10-750-185-56685/C
Sequence 56685, Application US/10750185
Publication No. US200502603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFIELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR IMPROVING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 56685
LENGTH: 1218
TYPE: DNA
ORGANISM: Bovine 1986680824166
US-10-750-185-56685
```

```
Query Match          4.2%; Score 34.6; DB 8; Length 1218;
Best Local Similarity 50.9%; Pred. No.2.9;
Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 631 TATGCCAAGTTTCACGACCAAAACAGCGATTGGGAGCGCTGATGCAACGACCGCCCTC 690
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 913 TAACCGAAGATTCAGCTTGAACAGAGCCCAAGTCAGCTTACCGGCAAGGCGCAGGTC 854
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 691 CTGAACCGCAACAGCAAGAGCGGATTTCCCGCCCAACATTTCAACACGCGCTGATC 750
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 853 CAGGCGTGCGCTGCGCAGTCGCTCCAGCGCGGCGCCGAACTGACAGGCGCGCTCCAG 794
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 751 CGGCGCTTGACCAAAATTCAGCGGGAAGGGAAGGAAAGGCGCG 791
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 793 GCCCCAGTCTCAAAAGGCTCCCGGCAAGCTTCCAACTCCG 753
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 9

```
US-10-750-623-56685/C
Sequence 56685, Application US/10750623
Publication No. US20050287531A1
GENERAL INFORMATION:
```

```
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFIELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR IMPROVING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 56685
LENGTH: 1218
TYPE: DNA
ORGANISM: Bovine 1986680824166
US-10-750-623-56685
```

```
Query Match          4.2%; Score 34.6; DB 8; Length 1218;
Best Local Similarity 50.9%; Pred. No.2.9;
Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 631 TATGCCAAGTTTCACGACCAAAACAGCGATTGGGAGCGCTGATGCAACGACCGCCCTC 690
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 913 TAACCGAAGATTCAGCTTGAACAGAGCCCAAGTCAGCTTACCGGCAAGGCGCAGGTC 854
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 691 CTGAACCGCAACAGCAAGAGCGGATTTCCCGCCCAACATTTCAACACGCGCTGATC 750
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 853 CAGGCGTGCGCTGCGCAGTCGCTCCAGCGCGGCGCCGAACTGACAGGCGCGCTCCAG 794
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 751 CGGCGCTTGACCAAAATTCAGCGGGAAGGGAAGGAAAGGCGCG 791
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 793 GCCCCAGTCTCAAAAGGCTCCCGGCAAGCTTCCAACTCCG 753
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 10

```
US-10-330-773-826/C
Sequence 826, Application US/10330773
Publication No. US20060040262A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc Malandro
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001300
CURRENT APPLICATION NUMBER: US/10/330,773
CURRENT FILING DATE: 2002-12-27
NUMBER OF SEQ ID NOS: 981
```


Db 191 TCAGAGGGCACTCTGGCCACACAGGAGGGCCGACCCCGGCTCTCCGATACCA 250
QY 473 TTCCCGAAGGAT 488
Db 251 GATGCTGGAATGTGT 266

RESULT 14

US-11-096-281-10/c
; Sequence 10, Application US/11096281
; Publication No. US20050267036A1
; GENERAL INFORMATION:
; APPLICANT: GARRY, MARY
; APPLICANT: BEZPROZVANNY, ILVA
; TITLE OF INVENTION: PEPTIDES OF CAV2.2 THAT INHIBIT PAIN
; FILE REFERENCE: UTSD:1558US
; CURRENT APPLICATION NUMBER: US/11/096,281
; PRIOR FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 60/558,383
; PRIOR FILING DATE: 2004-04-01
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 7364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (146)..(7165)
US-11-096-281-10

Query Match 4.1%; Score 33.6; DB 14; Length 7364;
Best Local Similarity 63.8%; Pred. No. 10;
Matches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 255 CGTATTGAGAGCAGCTTACTCGGCGAGAGGTGAGAAATTCCTTGCAGAGCGC 314
Db 407 CGTATTGAGAGCAGCTTACTCGGCGAGAGGTGAGAAATTCCTTGCAGAGCGC 348
QY 315 TTGGCTGCAAGAACGCTTG 334
Db 347 AGTTCTGCTTGACCGGAGTG 328

RESULT 15

US-09-925-065A-525973
; Sequence 525973, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 525973
; LENGTH: 626
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-525973

Query Match 4.0%; Score 33.4; DB 6; Length 626;
Best Local Similarity 47.8%; Pred. No. 5.6; Mismatches 106; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 344 CCGCTTTATTCGTCGCTTGAAACGATGTTATGACGCTCTGACCTCGCCTCGGCG 403
Db 26 CCGCTTTATTCGTCGCTTGAAACGATGTTATGACGCTCTGACCTCGCCTCGGCG 85
QY 404 TGGCGGATTACTGCGGCGCGCTTCCGCTGTGAAAGCGAACCTGGGGAGCGCG 463
Db 86 CTCGAGAACTGAGAGCAGTAGCTTCTGTGTTAAGTCAACCCAGTGTGTACTAT 145
QY 464 GCTATATCATTTCCGAAAGAGATGCGGTTTCTCTGACAGCTTTCGCGCCCGCG 523
Db 146 GCTATGCGAGCCCTAGAAATGATTCAGATTATCAACCTGTTGAGATGTCCAG 205
QY 524 CCGAAGGCTGCAACCCGTCGAT 546
Db 206 CCGTGTGGCAGAGCCAGGAGCT 228

Search completed: April 7, 2006, 21:23:57
Job time : 969.604 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:20:42 ; Search time 9845.41 Seconds
(without alignments)
4650.028 Million cell updates/sec

Title: US-09-211-691-1
Perfect score: 828
Sequence: 1 acgcaaacacgtatcacg.....tcattgcttcctccacataa 828

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 79147668 seqs, 27645789525 residues
Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/PCTUSA COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075 COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076 COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077 COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078 COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079 COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080 COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081 COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082 COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083 COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084 COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085 COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086 COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087 COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088 COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089 COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090 COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091 COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092 COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093 COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094 COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101 COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A COMB.seq:*

Result No.	Score	Query Match	Length	ID	Description
1	828	100.0	828	22	US-09-211-691-1
2	828	100.0	828	49	US-10-317-428-1
3	828	100.0	828	49	US-10-317-773-1
4	826.4	99.8	46594	33	US-09-806-866A-4
5	826.4	99.8	46594	63	US-10-915-740A-4
6	826.4	99.8	2242716	33	US-09-806-866A-1068
7	826.4	99.8	2242716	63	US-10-915-740A-1068
8	826.4	99.8	2272325	40	US-10-018-470A-1
9	826.4	99.8	5859	40	US-10-096-129-1
10	826.4	99.8	5859	54	US-10-654-528-1
11	826.4	99.8	5859	54	US-10-654-528-1
12	826.4	99.8	5859	70	US-11-102-497-1
13	826.4	99.8	5859	70	US-11-102-497-1
14	826.4	99.8	837	51	US-10-467-657-1309
15	826.4	99.8	837	51	US-10-467-657-1309
16	826.4	99.8	906	74	US-60-068-138-355
17	826.4	99.8	840	51	US-10-467-657-1325
18	826.4	99.8	840	51	US-10-467-657-1325
19	826.4	99.8	840	51	US-10-467-657-1325
20	826.4	99.8	840	51	US-10-467-657-1325
21	826.4	99.8	939	82	US-60-691-214-1378
22	826.4	99.8	14547	2	PCT-US04-17092-570
23	826.4	99.8	14547	61	PCT-US04-07001-570
24	826.4	99.8	14547	61	US-10-795-159-570

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```

QY 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCAATTGCCGAT 60
Db 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCAATTGCCGAT 60
QY 61 ACCTTGGCAGGACCGGCAATCCGTTGAGTTTTCAGGACTGATGCGCTGAAAG 120
Db 61 ACCTTGGCAGGACCGGCAATCCGTTGAGTTTTCAGGACTGATGCGCTGAAAG 120
QY 121 CTGGAACAGGCAATGAGGGAATCGTCCCGGCTGTCGCGGCAATTTGAGCGGA 180
Db 121 CTGGAACAGGCAATGAGGGAATCGTCCCGGCTGTCGCGGCAATTTGAGCGGA 180
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATTTGGAAGAGGAGGATTTGA 240
Db 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATTTGGAAGAGGAGGATTTGA 240
QY 241 CTGCGCTATATCAACGTAATTTGAGAGAGAGTTTACTCGCGGAGGATGAGAAA 300
Db 241 CTGCGCTATATCAACGTAATTTGAGAGAGAGTTTACTCGCGGAGGATGAGAAA 300
QY 301 CTGCGGAGAGAGCTGGGCTGCAAGAAAGCTTTGACCGGATACCGGCTTATCGTCCG 360
Db 301 CTGCGGAGAGAGCTGGGCTGCAAGAAAGCTTTGACCGGATACCGGCTTATCGTCCG 360
QY 361 TTGGAAGAGATGTTTATGACAGTCTGACCTCGGCTCGGCGTGGGAGATTTACTG 420
Db 361 TTGGAAGAGATGTTTATGACAGTCTGACCTCGGCTCGGCGTGGGAGATTTACTG 420
QY 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGAGGAGAGCGGAGGCTATATTCCTCGA 480
Db 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGAGGAGAGCGGAGGCTATATTCCTCGA 480
QY 481 AAAGGATGAGGTTTTCCTGCAAGGTTGCGGCTGCGCGGCAAGGCTGCAAGGCT 540
Db 481 AAAGGATGAGGTTTTCCTGCAAGGTTGCGGCTGCGCGGCAAGGCTGCAAGGCT 540
QY 541 GTGCAATGATGATGTTTATGAGAGGATTTTTCGACAGGAGAGGAGGATTTTCCAG 600
Db 541 GTGCAATGATGATGTTTATGAGAGGATTTTTCGACAGGAGAGGAGGATTTTCCAG 600
QY 601 AATCCGCGCTTGTGCGCCCAAGAGGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660
Db 601 AATCCGCGCTTGTGCGCCCAAGAGGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660
QY 661 TTGGGAGGCTGATTCGAACAGGCGGCTTCTGTAACCGCAAAACAGGAGCGGATTC 720
Db 661 TTGGGAGGCTGATTCGAACAGGCGGCTTCTGTAACCGCAAAACAGGAGCGGATTC 720
QY 721 CCGGCAACAGATTTCAAAACAGGCGGCTGATCCGCGCTTGAACAAATCAGAGGAAAG 780
Db 721 CCGGCAACAGATTTCAAAACAGGCGGCTGATCCGCGCTTGAACAAATCAGAGGAAAG 780
QY 781 GAAAAACGCGGCAAGAGGCGGCAAGGTTCAATTTGCTTTCCAAATTA 828
Db 781 GAAAAACGCGGCAAGAGGCGGCAAGGTTCAATTTGCTTTCCAAATTA 828

```

```

RESULT 3
US-10-317-773-1
; Sequence 1, Application US/10317773
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNac 4' Epimerase and a
; FILE REFERENCE: b19633-000812US
; CURRENT APPLICATION NUMBER: US/10/317,773
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15

```

```

; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(828)
; OTHER INFORMATION: beta-1,4-galactosyltransferase (1gfb)
; US-10-317-773-1
Query Match 100.0%; Score 828; DB 49; Length 828;
Best Local Similarity 100.0%; Pred. No. 5,6e-282;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCAATTGCCGAT 60
Db 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCAATTGCCGAT 60
QY 61 ACCTTGGCAGGACCGGCAATCCGTTGAGTTTTCAGGACTGATGCGCTGAAAG 120
Db 61 ACCTTGGCAGGACCGGCAATCCGTTGAGTTTTCAGGACTGATGCGCTGAAAG 120
QY 121 CTGGAACAGGCAATGAGGGAATCGTCCCGGCTGTCGCGGCAATTTGAGCGGA 180
Db 121 CTGGAACAGGCAATGAGGGAATCGTCCCGGCTGTCGCGGCAATTTGAGCGGA 180
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATTTGGAAGAGGAGGATTTGA 240
Db 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATTTGGAAGAGGAGGATTTGA 240
QY 241 CTGCGCTATATCAACGTAATTTGAGAGAGAGTTTACTCGCGGAGGATGAGAAA 300
Db 241 CTGCGCTATATCAACGTAATTTGAGAGAGAGTTTACTCGCGGAGGATGAGAAA 300
QY 301 CTGCGGAGAGAGCTGGGCTGCAAGAAAGCTTTGACCGGATACCGGCTTATCGTCCG 360
Db 301 CTGCGGAGAGAGCTGGGCTGCAAGAAAGCTTTGACCGGATACCGGCTTATCGTCCG 360
QY 361 TTGGAAGAGATGTTTATGACAGTCTGACCTCGGCTCGGCGTGGGAGATTTACTG 420
Db 361 TTGGAAGAGATGTTTATGACAGTCTGACCTCGGCTCGGCGTGGGAGATTTACTG 420
QY 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGAGGAGAGCGGAGGCTATATTCCTCGA 480
Db 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGAGGAGAGCGGAGGCTATATTCCTCGA 480
QY 481 AAAGGATGAGGTTTTCCTGCAAGGTTGCGGCTGCGCGGCAAGGCTGCAAGGCT 540
Db 481 AAAGGATGAGGTTTTCCTGCAAGGTTGCGGCTGCGCGGCAAGGCTGCAAGGCT 540
QY 541 GTGCAATGATGATGTTTATGAGAGGATTTTTCGACAGGAGAGGAGGATTTTCCAG 600
Db 541 GTGCAATGATGATGTTTATGAGAGGATTTTTCGACAGGAGAGGAGGATTTTCCAG 600
QY 601 AATCCGCGCTTGTGCGCCCAAGAGGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660
Db 601 AATCCGCGCTTGTGCGCCCAAGAGGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660
QY 661 TTGGGAGGCTGATTCGAACAGGCGGCTTCTGTAACCGCAAAACAGGAGCGGATTC 720
Db 661 TTGGGAGGCTGATTCGAACAGGCGGCTTCTGTAACCGCAAAACAGGAGCGGATTC 720
QY 721 CCGGCAACAGATTTCAAAACAGGCGGCTGATCCGCGCTTGAACAAATCAGAGGAAAG 780
Db 721 CCGGCAACAGATTTCAAAACAGGCGGCTGATCCGCGCTTGAACAAATCAGAGGAAAG 780
QY 781 GAAAAACGCGGCAAGAGGCGGCAAGGTTCAATTTGCTTTCCAAATTA 828
Db 781 GAAAAACGCGGCAAGAGGCGGCAAGGTTCAATTTGCTTTCCAAATTA 828

```

```
RESULT 4
US-09-806-866A-4/c
; Sequence 4, Application US/09806866A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteilin, Herre
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarcelllo, Maria
; APPLICANT: Scarcelllo, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: CHIR0313
; CURRENT APPLICATION NUMBER: US/09/806,866A
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 46594
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-09-806-866A-4

Query Match          99.8%; Score 826.4; DB 33; Length 46594;
Best Local Similarity 99.9%; Pred. No. 6.6e-280;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGCAAAACACGATTATAGCTTACGCTCCGCGCAAGACGAGCGGCACATTGCCAT 60
DB      33537  ATGCAAAACACGATTATAGCTTACGCTCCGCGCAAGACGAGCGGCACATTGCCAT 33478
QY      61  ACCCTCGGAGGACACGCGCATCCCGTTTCAGTTTTCAGCACTGATGCGCTGTAAGG 120
DB      33477  ACCCTCGGAGGACACGCGCATCCCGTTTCAGTTTTCAGCACTGATGCGCTGTAAGG 33418
QY      121  CTGGAACAGGCAATGCGGCACTCCGCTCCGCGCTGTGCGGCAACCCCTATTGAGCGGA 180
DB      33417  CTGGAACAGGCAATGCGGCACTCCGCTCCGCGCTGTGCGGCAACCCCTATTGAGCGGA 33358
QY      181  GTGGAAGAAAGCTGCTTTATGAGCCACGCGGATTTGTGAGACAGGCAATTGACGAAGGT 240
DB      33357  GTGGAAGAAAGCTGCTTTATGAGCCACGCGGATTTGTGAGACAGGCAATTGACGAAGGT 33298
QY      241  CTGCGGTATATACCGTATTTTGAAGACGATTTTACTGCGGAGAGGTGAGAAAAATTC 300
DB      33297  CTGCGGTATATACCGTATTTTGAAGACGATTTTACTGCGGAGAGGTGAGAAAAATTC 33238
QY      301  CTTGCGGAAGACGCTTGCGTGAAGAACGCTTTGACCCCGGATACCGCTTTATCGTCCGC 360
DB      33237  CTTGCGGAAGACGCTTGCGTGAAGAACGCTTTGACCCCGGATACCGCTTTATCGTCCGC 33178
QY      361  TTGGAAGACGATGTTTATGACAGCTGACCTGCGCCCTCGGGGTGGCGGATTAATGCGGG 420
DB      33177  TTGGAAGACGATGTTTATGACAGCTGACCTGCGCCCTCGGGGTGGCGGATTAATGCGGG 33118
QY      421  CGCGCTTTTCGCTGTTGGAAGACGAACCTGCGGGGACGCGCGGCTATATCATTTCCGA 480
DB      33117  CGCGCTTTTCGCTGTTGGAAGACGAACCTGCGGGGACGCGCGGCTATATCATTTCCGA 33058
```

```
QY      481  AAAGCATGCGGTTTTTCTTGACAGGTTTTCGCGCCCTGCGCCGCCGGAAGGCTGACCCC 540
DB      33057  AAAGCATGCGGTTTTTCTTGACAGGTTTTCGCGCCCTGCGCCGCCGGAAGGCTGACCCC 32998
QY      541  GTGCATCTGATGATGTTTATGAGGATTTTTCGACAGGAAGAAATGCGGTTTGCAGCTC 600
DB      32997  GTGCATCTGATGATGTTTATGAGGATTTTTCGACAGGAAGAAATGCGGTTTGCAGCTC 32938
QY      601  AATCCGCTTGTCGCGCCCAAGAGCTGATTAATGCAAGTTTACGACCAAAACAGCGCA 660
DB      32937  AATCCGCTTGTCGCGCCCAAGAGCTGATTAATGCAAGTTTACGACCAAAACAGCGCA 32878
QY      661  TTGGCAGCGTATGGAACAGACGCGCTCTGAAACCGCAACAGCAAGGCGCATTC 720
DB      32877  TTGGCAGCGTATGGAACAGACGCGCTCTGAAACCGCAACAGCAAGGCGCATTC 32818
QY      721  CCGGCAACACATTCAACACCGGCTGATTCGCGCTTGACCAAAATGACGAGGAAAG 780
DB      32817  CCGGCAACACATTCAACACCGGCTGATTCGCGCTTGACCAAAATGACGAGGAAAG 32758
QY      781  GAAAAACGCGGCAAGGCGGAAACAGTTCAATTGCTTCCATTA 828
DB      32757  GAAAAACGCGGCAAGGCGGAAACAGTTCAATTGCTTCCATTA 32710

RESULT 5
US-10-915-740A-4/c
; Sequence 4, Application US/10915740A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteilin, Herre
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarcelllo, Maria
; APPLICANT: Scarcelllo, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; PRIOR FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 46594
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-4

Query Match          99.8%; Score 826.4; DB 63; Length 46594;
Best Local Similarity 99.9%; Pred. No. 6.6e-280;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGCAAAACACGATTATAGCTTACGCTCCGCGCAAGACGAGCGGCACATTGCCAT 60
DB      33537  ATGCAAAACACGATTATAGCTTACGCTCCGCGCAAGACGAGCGGCACATTGCCAT 33478
QY      61  ACCCTCGGAGGACACGCGCATCCCGTTTCAGTTTTCAGCACTGATGCGCTGTAAGG 120
```

```

Db 33477 ACCCTGGCGAGGACCGGCATCCGTTTCAGTTTTCAGACGACGATGCGTCTGAAGG 33418
QY 121 CTGGAACAGGCATGAGGGAACCTGCTCCGCGCTGCGCGGACCCCTTTTGAAGGGA 180
Db 33417 CTGGAACAGGCATGAGGGAACCTGCTCCGCGCTGCGCGGACCCCTTTTGAAGGGA 33358
QY 181 GTGGAAGAAAGCTGCTTATGACGACGCGCTATTTGGAAGAGGATTTGAAGAGT 240
Db 33357 GTGGAAGAAAGCTGCTTATGACGACGCGCTATTTGGAAGAGGATTTGAAGAGT 33298
QY 241 CTGCGCATATTCACCGTATTTGAGAGAGAGTTTACTCGGCGAAGGTGAGAAATTC 300
Db 33297 CTGCGCATATTCACCGTATTTGAGAGAGAGTTTACTCGGCGAAGGTGAGAAATTC 33238
QY 301 CTGCGGAGAGAGCTGCTGCGGAGAGAGCTTTGACCCGATACCGCTTTATGTCGCC 360
Db 33237 CTGCGGAGAGAGCTGCTGCGGAGAGAGCTTTGACCCGATACCGCTTTATGTCGCC 33178
QY 361 TTGGAAGAGATGTTTATGACGCTGACCTGCGCCCTGCGCGGAGATTACTGCGGG 420
Db 33177 TTGGAAGAGATGTTTATGACGCTGACCTGCGCCCTGCGCGGAGATTACTGCGGG 33118
QY 421 CCGGCTTTCCGCTGTTGGAAGAGAGACCTGCGGAGAGCTATATCATTTTCCGA 480
Db 33117 CCGGCTTTCCGCTGTTGGAAGAGAGACCTGCGGAGAGCTATATCATTTTCCGA 33058
QY 481 AAGGCAATGAGTTTTCCTGACAGGTTTCCGCGCTGCGCGGAGAGGCTGACCCC 540
Db 33057 AAGGCAATGAGTTTTCCTGACAGGTTTCCGCGCTGCGCGGAGAGGCTGACCCC 32998
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACGAGAGAGAGTCCGAGCTC 600
Db 32997 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACGAGAGAGTCCGAGCTC 32938
QY 601 AATCCGCTTGTGCGCCGACGAGCTGATTTATGCAAGTTTTCAGACGACGAG 660
Db 32937 AATCCGCTTGTGCGCCGACGAGCTGATTTATGCAAGTTTTCAGACGAGAG 32878
QY 661 TTGGGAGAGCTGATGACAGACGCGCTCTGACACCGCAAGAGGCGGATTC 720
Db 32877 TTGGGAGAGCTGATGACAGACGCGCTCTGACACCGCAAGAGGCGGATTC 32818
QY 721 CCGGCAACATTTTAAACACGCGCTGATCCGCGCTTGAACCAATTCAGAGGAAAG 780
Db 32817 CCGGCAACATTTTAAACACGCGCTGATCCGCGCTTGAACCAATTCAGAGGAAAG 32758
QY 781 GAAAAAGCGCGGCAAGGCGGCAAGCTTATTTGCTCTTCCAAATTA 828
Db 32757 GAAAAAGCGCGGCAAGGCGGCAAGCTTATTTGCTCTTCCAAATTA 32710

```

```

RESULT 6
US-09-806-866A-1068/c
; Sequence 1068, Application US/09806866A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteijn, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Masimani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Rattai, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scariato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: CH100313
; CURRENT APPLICATION NUMBER: US/09/806,866A

```

```

; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: USN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1068
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-09-806-866A-1068

Query Match 99.8%; Score 826.4; DB 33; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 1,9e-278;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCAAAACAGGTTATGAGCTTACGCTTCGCGGAGAGAGGCGGACATTTGCCGAT 60
Db 1997047 ATGCAAAACAGGTTATGAGCTTACGCTTCGCGGAGAGAGGCGGACATTTGCCGAT 1996988
QY 61 ACCCTGGAGGACGCGCATCCGTTTCAGTTTTCAGAGCACTGATGCCGCTGAAAG 120
Db 1996987 ACCCTGGAGGACGCGCATCCGTTTCAGTTTTCAGAGCACTGATGCCGCTGAAAG 1996928
QY 121 CTGGAACAGGCAATGCGGAACTGCTCCGCGCTTGTGCGGACCCCTATTTGACGGA 180
Db 1996927 CTGGAACAGGCAATGCGGAACTGCTCCGCGCTTGTGCGGACCCCTATTTGACGGA 1996868
QY 181 GTGGAAGAAAGCTGCTTATGAGGAGCGCGCATTTGTGGAAGAGGCAATTTGAAGAGT 240
Db 1996867 GTGGAAGAAAGCTGCTTATGAGGAGCGCGCATTTGTGGAAGAGGCAATTTGAAGAGT 1996808
QY 241 CTGCGCATATACCGTATTTTGAAGAGAGCTTTTACTGCGGCAAGGTGAGAAATTC 300
Db 1996807 CTGCGCATATACCGTATTTTGAAGAGAGCTTTTACTGCGGCAAGGTGAGAAATTC 1996748
QY 301 CTGCGCAAGAGCGCTTGTGAGAGAGCGCTTGAACCGCGATACCGCTTTATGTCGCC 360
Db 1996747 CTGCGCAAGAGCGCTTGTGAGAGAGCGCTTGAACCGCGATACCGCTTTATGTCGCC 1996688
QY 361 TTGGAAGAGATGTTTATGACGCTGACCTGCGCTCCGCGGTGCGGATTAATGCGGG 420
Db 1996687 TTGGAAGAGATGTTTATGACGCTGACCTGCGCTCCGCGGTGCGGATTAATGCGGG 1996628
QY 421 CCGGCTTTCCGCTGTTGGAAGAGAGCACTGCGGAGAGCGGCGGCTATATCATTTCCGA 480
Db 1996627 CCGGCTTTCCGCTGTTGGAAGAGAGCACTGCGGAGAGCGGCGGCTATATCATTTCCGA 1996568
QY 481 AAGGCAATGAGTTTTCCTGACAGGTTTTCGCGCTGCGCGGAGAGGCTGACCCC 540
Db 1996567 AAGGCAATGAGTTTTCCTGACAGGTTTTCGCGCTGCGCGGAGAGGCTGACCCC 1996508
QY 541 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACGAGAGAGATGCGGTTTGCAGCTC 600
Db 1996507 GTGATCTGATGATGTTTTCAGCGATTTTTCAGACGAGAGAGATGCGGTTTGCAGCTC 1996448
QY 601 AATCCGCTTGTGCGCCGACGAGCTGATTTATGCAAGTTTTCAGACGACGAGAG 660
Db 1996447 AATCCGCTTGTGCGCCGACGAGCTGATTTATGCAAGTTTTCAGACGAGAGAG 1996388
QY 661 TTGGGAGAGCTGATGACAGACGCGCTCTGAAACCGCAACAGCAAGGCGGATTC 720
Db 1996387 TTGGGAGAGCTGATGACAGACGCGCTCTGAAACCGCAACAGCAAGGCGGATTC 1996328
QY 721 CCGGCAACATTTTAAACACGCGCTGATCCGCGCTTGAACCAATTCAGAGGAAAG 780
Db 1996327 CCGGCAACATTTTAAACACGCGCTGATCCGCGCTTGAACCAATTCAGAGGAAAG 1996268
QY 781 GAAAAAGCGCGGCAAGGCGGCAAGCTTATTTGCTCTTCCAAATTA 828

```

Db 1996267 GAAAAACCGCGCAAGGCGCGCAACGTTTCATTGCTTTCCATTA 1996220

RESULT 7

US-10-915-740A-1068/c
; Sequence 1068, Application US/10915740A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tettelin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroa
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scariato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; PRIOR FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 1068
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-1068

Query Match 99.8%; Score 826.4; DB 63; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 1.9e-278;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATCAAGCTTACGCTTCGCGCGCAAGACGAGCGCGCAATTGCGCAT 60
Db 1997047 ATGCAAAACCAAGTATCAAGCTTACGCTTCGCGCGCAAGACGAGCGCGCAATTGCGCAT 1996988
Qy 61 ACCTTGCGGAGGAGCGGCAATCCCGTTTTCAGTTTTCAGAGCACTGATGCGCTGAAAGG 120
Db 1996987 ACCTTGCGGAGGAGCGGCAATCCCGTTTTCAGTTTTCAGAGCACTGATGCGCTGAAAGG 1996928
Qy 121 CTGGAACAGGCAATGCGGGAATCGTCCCGGCTTGTGCGCGCACTCCCTAATTGAGCGGA 180
Db 1996927 CTGGAACAGGCAATGCGGGAATCGTCCCGGCTTGTGCGCGCACTCCCTAATTGAGCGGA 1996868
Qy 181 GTGGAAGAAAGCGCTTATGAGCGACGCGGATTTGAGAGGAGGATTTGAGCAAGGT 240
Db 1996867 GTGGAAGAAAGCGCTTATGAGCGACGCGGATTTGAGAGGAGGATTTGAGCAAGGT 1996808
Qy 241 CTGCGGTATATCAAGCTTATTTGAGAGCGAGCTTTTACTGCGGAGAGGTGAGAAAATTTC 300
Db 1996807 CTGCGGTATATCAAGCTTATTTGAGAGCGAGCTTTTACTGCGGAGAGGTGAGAAAATTTC 1996748
Qy 301 CTGCGCGAAGAGCGCTTGTGAGCAAGACCTTTGACCCGGAATACGCTTTATCTGCGCG 360
Db 1996747 CTGCGCGAAGAGCGCTTGTGAGCAAGACCTTTGACCCGGAATACGCTTTATCTGCGCG 1996688
Qy 361 TTGGAACGATGTTTATGACAGCTTCTGAGCTCGCCCTCCGCGGTGGCGGATTAATGCGGG 420
Db 1996687 TTGGAACGATGTTTATGACAGCTTCTGAGCTCGCCCTCCGCGGTGGCGGATTAATGCGGG 1996628

Qy 421 CGCGCTTTCCGCTTTGAAAAGCGAACAATGCGGAGGAGCGGCGCTATATCAATTTCCGGA 480
Db 1996627 CGCGCTTTCCGCTTTGAAAAGCGAACAATGCGGAGGAGCGGCGCTATATCAATTTCCGGA 1996568
Qy 481 AAAGCATGCGGTTTTCCTGACAGTTTTCGCGCTGCGGCCGGAAGGCTGCAACCC 540
Db 1996567 AAAGCATGCGGTTTTCCTGACAGTTTTCGCGCTGCGGCCGGAAGGCTGCAACCC 1996508
Qy 541 GTGATCTGATGATGTTTTCAGGATTTTTCAGACAGGAGAAATGCGGTTTCCAGTTC 600
Db 1996507 GTGATCTGATGATGTTTTCAGGATTTTTCAGACAGGAGAAATGCGGTTTCCAGTTC 1996448
Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAACGACCAAAACAGCGCA 660
Db 1996447 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAACGACCAAAACAGCGCA 1996388
Qy 661 TTGGGACGCTGATGAAACAGACCGCTCTGAAACCGCAAAAGGCGGATTC 720
Db 1996387 TTGGGACGCTGATGAAACAGACCGCTCTGAAACCGCAAAAGGCGGATTC 1996328
Qy 721 CCCGCAACACATTTAAACCGCGCTGATCCGCGCTTGAACCAAAATGACAGGAGAAAG 780
Db 1996327 CCCGCAACACATTTAAACCGCGCTGATCCGCGCTTGAACCAAAATGACAGGAGAAAG 1996268
Qy 781 GAAAAACCGCGCAAGGCGGGAACGTTCAATGTCCTTCCATTA 828
Db 1996267 GAAAAACCGCGCAAGGCGGGAACGTTCAATGTCCTTCCATTA 1996220

RESULT 8

US-10-018-470A-1/c
; Sequence 1, Application US/10018470A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tettelin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroa
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scariato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Research
; FILE REFERENCE: CHIR-0319
; CURRENT APPLICATION NUMBER: US/10/018,470A
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: GB-0004695.3
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/US/05928
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 2272325
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-018-470A-1

Query Match 99.8%; Score 826.4; DB 40; Length 2272325;
Best Local Similarity 99.9%; Pred. No. 1.9e-278;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATCAAGCTTACGCTTCGCGCGAAGACGAGGCGCAATTCGCGAT 60
|||||


```

Db 202656 ATGCAAAACCGATTACGCTTAGCTTCGCGCAGAGGAGGCGACATTGCCGAT 2026597
Qy 61 ACCCTGGGAGGAGCGGAGCATCCGTTTCAGTTTTTGACGACGATGATCCGCTGAAAG 120
Db 2026596 ACCCTGGGAGGAGCGGAGCATCCGTTTCAGTTTTTGACGACGATGATCCGCTGAAAG 2026537
Qy 121 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCGACCCCTATTGAGCGGA 180
Db 2026536 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCGACCCCTATTGAGCGGA 2026477
Qy 181 GTGGAAGAAAGCTGCTTTATGAGCAGCGGCTATTTGGAAGCAGGCAATGGAAGGT 240
Db 2026476 GTGGAAGAAAGCTGCTTTATGAGCAGCGGCTATTTGGAAGCAGGCAATGGAAGGT 2026417
Qy 241 CTGCGGTATATCAACCGTATTTGAGAGCAGAGGTTTACTGCGGAGAGGTGAGAAATTC 300
Db 2026416 CTGCGGTATATCAACCGTATTTGAGAGCAGAGGTTTACTGCGGAGAGGTGAGAAATTC 2026357
Qy 301 CTTCGGAAGACGCTTGGCTGCAAGACGCTTTGACCCGAGTACCGCTTTATCTGTCGC 360
Db 2026356 CTTCGGAAGACGCTTGGCTGCAAGACGCTTTGACCCGAGTACCGCTTTATCTGTCGC 2026297
Qy 361 TTGGAAGACGATGTTTATGACAGCTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGGG 420
Db 2026296 TTGGAAGACGATGTTTATGACAGCTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGGG 2026237
Qy 421 CGGCGCTTTCGCGCTTTGGAAGACGACGCTGGGAGACGGCGGCTATATCTTCCCGA 480
Db 2026236 CGGCGCTTTCGCGCTTTGGAAGACGACGCTGGGAGACGGCGGCTATATCTTCCCGA 2026177
Qy 481 AAGGGAATGCGATTTTTCCTGCAAGGTTTCCGCTCGCGCCGAGAGGCTGCAAGCC 540
Db 2026176 AAGGGAATGCGATTTTTCCTGCAAGGTTTCCGCTCGCGCCGAGAGGCTGCAAGCC 2026117
Qy 541 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 600
Db 2026116 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 2026057
Qy 601 AATCCGCGCTTGTGCGCCCAAGAGCTGACATTAAGCCAGTTTACAGCAAAACAGCGCA 660
Db 2026056 AATCCGCGCTTGTGCGCCCAAGAGCTGACATTAAGCCAGTTTACAGCAAAACAGCGCA 2025997
Qy 661 TTGGGAGGCTGATTCGACAGACGCGCTCTCTGAAACCGCAACAGAAAGGCGGATTC 720
Db 2025996 TTGGGAGGCTGATTCGACAGACGCGCTCTCTGAAACCGCAACAGAAAGGCGGATTC 2025937
Qy 721 CCGGCAACACATTGAAACACGCGCTGATTCGCGGCTTGAACAAATATGACAGGAAAG 780
Db 2025936 CCGGCAACACATTGAAACACGCGCTGATTCGCGGCTTGAACAAATATGACAGGAAAG 2025877
Qy 781 GAAAAACGCGGCAAGGCGGCAACGTTTCATGTTGCTTTTCCAAATTA 828
Db 2025876 GAAAAACGCGGCAAGGCGGCAACGTTTCATGTTGCTTTTCCAAATTA 2025829

```

```

RESULT 9
US-10-096-129-1
; Sequence 1, Application US/10096129
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: ROTH, STEPHEN
; APPLICANT: BUCZALA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO SACCHARIDE UNITS WITH A
; FILE REFERENCE: POLYGLYCOSYLTRANSFERASE
; CURRENT APPLICATION NUMBER: US/10/096,129
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/338,943
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 08/478,140
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-096-129-1
Query Match 87.7%; Score 726; DB 40; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6,2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 1 ATGCAAAACCGATTACGCTTAGCTTCGCGCAGAGGAGGCGACATTGCCGAT 60
Db 1491 ATGCAAAACCGATTACGCTTAGCTTCGCGCAGAGGAGGCGACATTGCCGCA 1550
Qy 61 ACCCTGGGAGGAGCGGAGCATCCGTTTCAGTTTTTGACGACGATGATCCGCTGAAAG 120
Db 1551 ACCCTGGGAGGAGCGGAGCATCCGTTTCAGTTTTTGACGACGATGATCCGCTGAAAG 1610
Qy 121 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCGACCCCTATTGAGCGGA 180
Db 1611 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCGACCCCTATTGAGCGGA 1670
Qy 181 GTGGAAGAAAGCTGCTTTATGAGCAGCGGCTATTTGGAAGCAGGCAATGGAAGGT 240
Db 1671 GTGGAAGAAAGCTGCTTTATGAGCAGCGGCTATTTGGAAGCAGGCAATGGAAGGT 1730
Qy 241 CTGCGGTATATCAACCGTATTTGAGAGCAGTTTTACTGCGGAGAGGTGAGAAATTC 300
Db 1731 CTGCGGTATATCAACCGTATTTGAGAGCAGTTTTACTGCGGAGAGGTGAGAAATTC 1790
Qy 301 CTTCGGAAGACGCTTGGCTGCAAGACGCTTTGACCCGAGTACCGCTTTATCTGTCGC 360
Db 1791 CTTCGGAAGACGCTTGGCTGCAAGACGCTTTGACCCGAGTACCGCTTTATCTGTCGC 1850
Qy 361 TTGGAAGACGATGTTTATGACAGCTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGGG 420
Db 1851 TTGGAAGACGATGTTTATGACAGCTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGGG 1910
Qy 421 CGGCGCTTTCGCGCTTTGGAAGACGACGCTGGGAGACGGCGGCTATATCTTCCCGA 480
Db 1911 CGGCGCTTTCGCGCTTTGGAAGACGACGCTGGGAGACGGCGGCTATATCTTCCCGA 1970
Qy 481 AAGGGAATGCGATTTTTCCTGCAAGGTTTCCGCTCGCGCCGAGAGGCTGCAAGCC 540
Db 1971 AAGGGAATGCGATTTTTCCTGCAAGGTTTCCGCTCGCGCCGAGAGGCTGCAAGCC 2030
Qy 541 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 600
Db 2031 GTGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 2090
Qy 601 AATCCGCGCTTGTGCGCCCAAGAGCTGACATTAAGCCAGTTTACAGCAAAACAGCGCA 660
Db 2091 AATCCGCGCTTGTGCGCCCAAGAGCTGACATTAAGCCAGTTTACAGCAAAACAGCGCA 2150
Qy 661 TTGGGAGGCTGATTCGACAGACGCGCTCTCTGAAACCGCAACAGAAAGGCGGATTC 720
Db 2151 TTGGGAGGCTGATTCGACAGACGCGCTCTCTGAAACCGCAACAGAAAGGCGGATTC 2210
Qy 721 CCGGCAACACATTGAAACACGCGCTGATTCGCGGCTTGAACAAATATGACAGGAAAG 780
Db 2211 CCGGCAACACATTGAAACACGCGCTGATTCGCGGCTTGAACAAATATGACAGGAAAG 2270
Qy 781 GAAAAACGCGGCAAGGCGGCAACGTTTCATGTTGCTTTTCCAAATTA 814
Db 2271 GAAAAACGCGGCAAGGCGGCAACGTTTCATGTTGCTTTTCCAAATTA 2304

```

```

RESULT 10
US-10-654-528-1
; Sequence 1, Application US/10654528
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND

```

```

; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(381)
; OTHER INFORMATION: gly (glycyl tRNA synthetase beta chain)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (445)..(1491)
; OTHER INFORMATION: 19tC
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2342)..(3262)
; OTHER INFORMATION: 19tC
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3322)..(4335)
; OTHER INFORMATION: 19tD
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4354)..(5196)
; OTHER INFORMATION: 19tE
; US-10-654-528-1

Query Match      87.7%; Score 726; DB 54; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6,2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```

```

; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 7
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1491)..(2330)
; OTHER INFORMATION: 19tB
; US-10-654-528-7

Query Match      87.7%; Score 726; DB 54; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6,2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```



```
QY 121 CTGAAACAGGCAATGGCGAACTGTCCTCCGCGCTTGTGCGGGAACCCCTATTGAGCGGA 180
DB 1611 CTGGAACGGGCAATGGCGAACTGTCCTCCGCGCTTGTGCGGGAACCCCTATTGAGCGGA 1670
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATGTGGAAGCAAGGATGGAAGT 240
DB 1671 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATGTGGAAGCAAGGATGGAAGT 1730
QY 241 CTGCGGATATGACCGTATTTGAGAGCAAGTTTACTCGGCGAAGGTGAGAAAATTC 300
DB 1731 GTACCGTATATCCGCTATTTGAGAGCAAGTTTACTCGGCGAAGGTGAGAGATTC 1790
QY 301 CTGCGGAAAGACGCTTGGCTGCAAGAACGCTTGAACCGGATACCGCTTTATCTGTCGC 360
DB 1791 CTGCGGAAAGATCTTGGCTGCAAGAACGCTTGAACCGGATACCGCTTTATCTGTCGC 1850
QY 361 TTGGAAGACGATGTTATGACAGTCTGCAAGCTCCGCGCTGCGGATTACTGCGGG 420
DB 1851 TTGGAAGACGATGTTATGACAGTCTGCAAGCTCCGCGCTGCGGATTACTGCGGG 1910
QY 421 CGGCGCTTTCGCTGTTGGAAGGCAACCTGCGGAGCGCGGCTATATCATTTCCGA 480
DB 1911 CGGCGCTTTCGCTGTTGGAAGGCAACCTGCGGAGCGCGGCTATATTTCCGA 1970
QY 481 AAGGCAATGGGCTTTTCTGCAAGGTTTGGCGCGCGCGCGCAAGGCTGCAACCC 540
DB 1971 AAGGCAATGGGCTTTTCTGCAAGGTTTGGCGCGCGCGCGCAAGGCTGCAACCC 2030
QY 541 GTGATCTGATGATGTTGTCAGCGATTTTTCGACAGGGAAGAAATGCGGTTTCCAGCTC 600
DB 2031 GTGATTTGATGATGTTTCCGCAACCTGACAGGGAAGAAATGCGGTTTCCAGCTC 2090
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGGCA 660
DB 2091 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGGCA 2150
QY 661 TTGGGAGCGCTGATCGAAGCAAGCGCTCTGTAACCGGAAACAGAAAGCGGATTC 720
DB 2151 TTGGGAGCGCTGATCGAAGCAAGCGCTCTGTAACCGGAAACAGAAAGCGGATTC 2210
QY 721 CCGGCAACATTCATTAACACCGCTGATCGGCGCTTGAACAAATCAGCAGGAAAG 780
DB 2211 CCGGCAACATTCATTAACACCGCTGATCGGCGCTTGAACAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGGCAAGAGCGCAACAGTTGATG 814
DB 2271 GAAAAACGCGGCAAGAGCGCAACAGTTGATG 2304

RESULT 12
US-11-102-497-1
; Sequence 1, Application US/1102497
; GENERAL INFORMATION:
; APPLICANT: Gofeschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
```

```
SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(381)
; OTHER INFORMATION: glys [glycy] tRNA synthetase beta chain)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (445)..(1491)
; OTHER INFORMATION: lgta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2342)..(3262)
; OTHER INFORMATION: lgta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3322)..(4335)
; OTHER INFORMATION: lgta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4354)..(5196)
; OTHER INFORMATION: lgta
US-11-102-497-1

Query Match 87.7%; Score 726; DB 70; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6,2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACAGCTATACGCTTACGCTTCCGCGCAAGAACGAGGCGCAATTCGCGAT 60
DB 1491 ATGCAAAACAGCTATACGCTTACGCTTCCGCGCAAGAACGAGGCGCAATTCGCGCA 1550
QY 61 ACCTGGGAGGCAAGCGCATCCCGTTTCAAGTTTTCGACGCACTGATGCGGTGAAAG 120
DB 1551 ACCTGGGAGGCAAGCGCATCCCGTTTCAAGTTTTCGACGCACTGATGCGGTGAAAG 1610
QY 121 CTGGAACAGGCAATGGCGGAACCTGTCGCGCTTGTGCGGCAACCCCTATTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGGCGGAACCTGTCGCGCTTGTGCGGCAACCCCTATTGAGCGGA 1670
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATGTGGAAGCAAGGATGGAAGT 240
DB 1671 GTGGAAGAAAGCCTGCTTATGAGCAAGCGCTATGTGGAAGCAAGGATGGAAGT 1730
QY 241 CTGCGGATATGACCGTATTTGAGAGCAAGTTTACTCGGCGAAGGTGAGAAAATTC 300
DB 1731 GTACCGTATATCCGCTATTTGAGAGCAAGTTTACTCGGCGAAGGTGAGAGATTC 1790
QY 301 CTGCGGAAAGACGCTTGGCTGCAAGAACGCTTGAACCGGATACCGCTTTATCTGTCGC 360
DB 1791 CTGCGGAAAGATCTTGGCTGCAAGAACGCTTGAACCGGATACCGCTTTATCTGTCGC 1850
QY 361 TTGGAAGACGATGTTATGACAGTCTGCAAGCTCCGCGCTGCGGATTACTGCGGG 420
DB 1851 TTGGAAGACGATGTTATGACAGTCTGCAAGCTCCGCGCTGCGGATTACTGCGGG 1910
QY 421 CGGCGCTTTCGCTGTTGGAAGGCAACCTGCGGAGCGCGGCTATATCATTTCCGA 480
DB 1911 CGGCGCTTTCGCTGTTGGAAGGCAACCTGCGGAGCGCGGCTATATTTCCGA 1970
QY 481 AAGGCAATGGGCTTTTCTGCAAGGTTTGGCGCGCGCGCAAGGCTGCAACCC 540
DB 1971 AAGGCAATGGGCTTTTCTGCAAGGTTTGGCGCGCGCGCAAGGCTGCAACCC 2030
QY 541 GTGATCTGATGATGTTGTCAGCGATTTTTCGACAGGGAAGAAATGCGGTTTCCAGCTC 600
DB 2031 GTGATTTGATGATGTTTCCGCAACCTGACAGGGAAGAAATGCGGTTTCCAGCTC 2090
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGGCA 660
```

Db 2091 AATCCGCGCTTGCGCCCAAGAGCTGATTATGCAAGTTTCAGACCAAAAACAGCGCA 2150
Qy 661 TTGGGACGCTGATATGGAACAGACCGCTCTCTGAACCGGAACAGCAAGGCGCATTC 720
Db 2151 TTGGGACGCTGATATGGAACAGACCGCTCTCTGAACCGGAACAGCAAGGCGCATTC 2210
Qy 721 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
Db 2211 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 2270
Qy 781 GAAAAACCGCGCAAGGCGGCAAGATTCATTG 814
Db 2271 GAAAAACCGCGCAAGGCGGCAAGATTCATTG 2304

RESULT 13
US-11-102-497-7
; Sequence 7, Application US/11102497
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; PRIOR FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 7
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1491)..(2330)
; OTHER INFORMATION: 1gCB
US-11-102-497-7

Query Match 87.7%; Score 726; DB 70; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6,2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATGAGCTTTCGCGCGAGAACGAGGCGGACATTCGCCAT 60
Db 1491 ATGCAAAACCAAGTATGAGCTTTCGCGCGAGAACGAGGCGGACATTCGCCCA 1550
Qy 61 ACCTTCGAGGACGAGCATCCCGTTTTCAGATTTTTCAGACGACTGATGCGCTGAAAG 120
Db 1551 ACCTTCGAGGACGAGCATCCCGTTTTCAGATTTTTCAGACGACTGATGCGCTGAAAG 1610
Qy 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTCGCGCACCCCTATTGTAGCGGA 180
Db 1611 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTCGCGCACCCCTATTGTAGCGGA 1670
Qy 181 GTGGAAAAAGCTGCTTTATGAGCCAGCGCGTATTGTGGAAGCAGGCAATTGACGAAGT 240
Db 1671 GTGGAAAAAGCTGCTTTATGAGCCAGCGCGTATTGTGGAAGCAGGCAATTGACGAAGT 1730
Qy 241 CTGCGGTATATACCGTATTGTGAGAGCAGTTTTATCTGCGGAGAGGTGAGGAAAAATTC 300
Db 1731 GTACGCTATATCGCGGTATTGTGGAAGATGATGCTTTATCTGCGGAGAGGTGAGGAGCATTC 1790

Qy 301 CTTCGCGAAGACGCTTGCTGCAAGAACGCTTTGACCCGAGATACCGCTTTATGCTCCGC 360
Db 1791 CTTCGCGAAGATCTTGCTGCTGCAAGAACGCTTTGACCCGAGATACCGCTTTATGCTCCGC 1850
Qy 361 TTGAAACGATGTTTATGACAGCTCTGATCCGCTTCGCGGCTGAGGCAATTACTGCGGG 420
Db 1851 TTGAAACGATGTTTATGACAGCTCTGATCCGCTTCGCGGCTGAGGCAATTACTGCGGG 1910
Qy 421 CGCGCTTTCGCGCTGTTGGAAGGAAACAAGTGGGGAAGCGGCGGTATATGATTTCCCGA 480
Db 1911 CGCGCTTTCGCGCTTTCGGAAGGAAACAAGTGGGGAAGCGGCGGTATATGATTTCCCGA 1970
Qy 481 AAAGGATGCGGTTTTCCTGGAAGGTTTCCCGCTGCGCGCGAAAGGCTGACATCC 540
Db 1971 AAAGGATGCGGTTTTCCTGGAAGGTTTCCCGCTGCGCGCGAAAGGCTGACATCC 2030
Qy 541 GTGATCTGATGATGTTTTCAGGATTTTTCGACAGGGAAGGAAATGCGGTTTTCAGGCTC 600
Db 2031 GTGATCTGATGATGTTTTCAGGATTTTTCGACAGGGAAGGAAATGCGGTTTTCAGGCTC 2090
Qy 601 AATCCGCGCTTGCGCCCAAGAGCTGATATGCAAGTTTCAGACCAACCAACAGCGCA 660
Db 2091 AATCCGCGCTTGCGCCCAAGAGCTGATATGCAAGTTTCAGACCAACCAACAGCGCA 2150
Qy 661 TTGGGACGCTGATATGGAACAGACCGCTCTCTGAACCGGAACAGCAAGGCGCATTC 720
Db 2151 TTGGGACGCTGATATGGAACAGACCGCTCTCTGAACCGGAACAGCAAGGCGCATTC 2210
Qy 721 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
Db 2211 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 2270
Qy 781 GAAAAACCGCGCAAGGCGGCAAGATTCATTG 814
Db 2271 GAAAAACCGCGCAAGGCGGCAAGATTCATTG 2304

RESULT 14
US-10-467-657-1309
; Sequence 1309, Application US/10467657
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 1309
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1309

Query Match 87.1%; Score 721.2; DB 51; Length 837;
Best Local Similarity 92.9%; Pred. No. 6e-244;
Matches 756; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATGAGCTTTCGCGCGAGAACGAGGCGGACATTCGCCAT 60
Db 1 ATGCAAAACCAAGTATGAGCTTTCGCGCGAGAACGAGGCGGACATTCGCCCA 60
Qy 61 ACCTTCGAGGACGAGCATCCCGTTTTCAGATTTTTCAGACGACTGATGCGCTGAAAG 120
Db 61 ACCTTCGAGGACGAGCATCCCGTTTTCAGATTTTTCAGACGACTGATGCGCTGAAAG 120
Qy 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTCGCGCACCCCTATTGTAGCGGA 180

```

Db      121 CTGAAACAGCGCATGCGGAACCTGTCCTCCGCTGTCGCGCACTCTATTTGAGCGGA 180
QY      181 GTGGAAGAAACCGCTTTATGAGCAACGCGGTATTTGGAAGCAGGATTTGACGAAGT 240
Db      181 GTGGAAGAAACCGCTTTATGAGCAACGCGGTATTTGGAAGCAGGATTTGAGT 240
QY      241 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTACTCGCGAGAGGTGAGAAATTC 300
Db      241 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTACTCGCGAGAGGTGAGAGTTC 300
QY      301 CTGCGGAAAGACGCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGTCGC 360
Db      301 CTGCGGAAAGATCTTGCTTGGAAGAGCGTTTGTATAGATTCCTGCTTTATCTGTCGC 360
QY      361 TTGGAAGACATGTTTATGACAGCTCTGACCTCGCCCTCGCGCTGCGGATTACTGCGG 420
Db      361 TTGGAAGACATGTTTATGACAGCTCTGACCTCGCCCTCGCGCTGCGGATTAAGCGG 420
QY      421 CGGCGCTTTCGCTGTTGGAAGCAGAACATGCGGAGGAGCGGCGGCTATATCATTTCCGA 480
Db      421 CGGCGCTTTCGCTGTTGGAAGCAGAACATGCGGAGGAGCGGCGGCTATATTTCCGA 480
QY      481 AAGCGATGCGGCTTTTCTGCAAGAGTTGCGCCCTGCGCGCGAGAGGCTGCAACCC 540
Db      481 AAGCGATGCGGCTTTTCTGCAAGAGTTGCGCCCTGCGCGCGAGAGGCTGCAACCC 540
QY      541 GTGATCTGATGATGTTTACAGCGATTTTTCGACAGGAGAGAAATGCGGTTTCCAGCTC 600
Db      541 GTGATCTGATGATGTTTACAGCGATTTTTCGACAGGAGAGAAATGCGGTTTCCAGCTC 600
QY      601 AATCCGCGCTTGCGCGCCCAAGAGCTGATATGCAAGTTTACAGACCAAAACAGCGCA 660
Db      601 AATCCGCGCTTGCGCGCCCAAGAGCTGATATGCAAGTTTACAGACCAAAACAGCGCA 660
QY      661 TTGCGAGCGCTGATGCAACAGACGCTCTGCAACGCGCAACAGCAAGGCGCATTC 720
Db      661 TTGCGAGCGCTGATGCAACAGACGCTCTGCAACGCGCAACAGCAAGGCGCATTC 720
QY      721 CCGGCAACACATTTCAACACGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db      721 CCGGCAACACATTTCAACACGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 780
QY      781 GAAAAACGCGGCAAGGCGCGCAAGCTTATG 814
Db      781 GAAAAACGCGGCAAGGCGCGCAAGCTATG 814

RESULT 15
US-10-467-657A-1309
/ Sequence 1309, Application US/10467657A
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE: 002441.00061
/ CURRENT APPLICATION NUMBER: US/10/467,657A
/ PRIOR FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqMan99, version 1.04
/ SEQ ID NO 1309
/ TYPE: DNA
/ LENGTH: 837
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657A-1309

Query Match      87.1%; Score 721.2; DB 51; Length 837;
Best Local Similarity 92.9%; Pred. No. 6e-244;
```

```

Matches 756; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY      1 ATGCAAAACCAAGTTATGAGCTTATGCTTCCGCGCAAGACGAGGCGCATTTGCCAT 60
Db      1 ATGCAAAACCAAGTTATGAGCTTATGCTTCCGCGCAAGACGAGGCGCATTTGCCA 60
QY      61 ACCCTGGAGAGGACGCGCATCCCGTTTCAAGTTTTCAGAGCATGATGCGGTGAAAG 120
Db      61 ACCCTGGAGAGGACGCGCATCCCGTTTCAAGTTTTCAGAGCATGATGCGGTGAAAG 120
QY      121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCTTGCGCGCAACCTTATTTGAGCGGA 180
Db      121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCTTGCGCGCAACCTTATTTGAGCGGA 180
QY      181 GTGGAAGAAACCGCTTTATGAGCAACGCGGTATTTGGAAGCAGGATTTGAGAGGT 240
Db      181 GTGGAAGAAACCGCTTTATGAGCAACGCGGTATTTGGAAGCAGGATTTGAGAGGT 240
QY      241 CTGCGGTATATCAACCGTATTTGAGAGCAGGTTTACTCGCGAGAGGTGAGAAATTC 300
Db      241 CTGCGGTATATCAACCGTATTTGAGAGCAGGTTTACTCGCGAGAGGTGAGAGTTC 300
QY      301 CTGCGGAAAGACGCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGTCGC 360
Db      301 CTGCGGAAAGATCTTGCTTGGAAGAGCGTTTGTATAGATTCCTGCTTTATCTGTCGC 360
QY      361 TTGGAAGACATGTTTATGACAGCTCTGACCTCGCCCTCGCGCTGCGGATTACTGCGG 420
Db      361 TTGGAAGACATGTTTATGACAGCTCTGACCTCGCCCTCGCGCTGCGGATTAAGCGG 420
QY      421 CGGCGCTTTCGCTGTTGGAAGCAGAACATGCGGAGGAGCGGCGGCTATATCATTTCCGA 480
Db      421 CGGCGCTTTCGCTGTTGGAAGCAGAACATGCGGAGGAGCGGCGGCTATATTTCCGA 480
QY      481 AAGCGATGCGGCTTTTCTGCAAGAGTTGCGCCCTGCGCGCGAGAGGCTGCAACCC 540
Db      481 AAGCGATGCGGCTTTTCTGCAAGAGTTGCGCCCTGCGCGCGAGAGGCTGCAACCC 540
QY      541 GTGATCTGATGATGTTTACAGCGATTTTTCGACAGGAGAGAAATGCGGTTTCCAGCTC 600
Db      541 GTGATCTGATGATGTTTACAGCGATTTTTCGACAGGAGAGAAATGCGGTTTCCAGCTC 600
QY      601 AATCCGCGCTTGCGCGCCCAAGAGCTGATATGCAAGTTTACAGACCAAAACAGCGCA 660
Db      601 AATCCGCGCTTGCGCGCCCAAGAGCTGATATGCAAGTTTACAGACCAAAACAGCGCA 660
QY      661 TTGCGAGCGCTGATGCAACAGACGCTCTGCAACGCGCAACAGCAAGGCGCATTC 720
Db      661 TTGCGAGCGCTGATGCAACAGACGCTCTGCAACGCGCAACAGCAAGGCGCATTC 720
QY      721 CCGGCAACACATTTCAACACGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 780
Db      721 CCGGCAACACATTTCAACACGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 780
QY      781 GAAAAACGCGGCAAGGCGCGCAAGCTTATG 814
Db      781 GAAAAACGCGGCAAGGCGCGCAAGCTATG 814

Search completed: April 7, 2006, 21:06:22
Job time : 9854.41 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 17:01:35 ; Search time 3557.5 Seconds
(without alignments)
706.685 Million cell updates/sec

Title: US-09-211-691-1
Perfect score: 828
Sequence: 1 atgcaaacacgttatcag.....tcattgcttcacataa 828

Scoring table: IDENTITY NUC
Gapex 10.0, Gapext 1.0

Searched: 6718506 seqs, 1518135245 residues

Total number of hits satisfying chosen parameters: 13437012

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents NA New: *
1: /SID55/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /SID55/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /SID55/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /SID55/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /SID55/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /SID55/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
8: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
9: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
10: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
11: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
12: /SID55/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
C 1	826.4	99.8	2242716	7 US-10-018-470B-1	Sequence 1, Appl1
C 2	36.4	4.4	2242716	7 US-10-018-470B-1	Sequence 1, Appl1
C 3	36.4	4.3	270	8 US-11-329-388-18070	Sequence 18070, A
C 4	36.4	4.3	286	10 US-11-266-748A-12825	Sequence 12825, A
C 5	36.4	4.3	286	10 US-11-266-748A-64567	Sequence 64567, A
C 6	36.4	4.3	286	10 US-11-266-748A-67339	Sequence 67339, A
C 7	36.4	4.3	1000	10 US-11-266-748A-11615	Sequence 11615, A
C 8	36.4	4.3	1000	10 US-11-266-748A-158329	Sequence 158329, A
C 9	36.4	4.3	1000	10 US-11-266-748A-286336	Sequence 286336, A
C 10	36.4	4.3	1000	10 US-11-266-748A-337765	Sequence 337765, A
C 11	36.4	4.3	1000	10 US-11-266-748A-396766	Sequence 396766, A
C 12	36.4	4.3	1000	10 US-11-266-748A-467812	Sequence 467812, A
C 13	36.4	4.3	2821	8 US-11-056-355B-89151	Sequence 89151, A
C 14	36.4	4.3	2821	8 US-11-056-355B-92907	Sequence 92907, A
C 15	35.6	4.3	516	8 US-11-330-364-14690	Sequence 14690, A
C 16	35.6	4.2	1241	10 US-11-266-748A-357498	Sequence 357498, A
C 17	35.6	4.2	1241	10 US-11-266-748A-440877	Sequence 440877, A
C 18	35.6	4.2	1704	7 US-10-531-147-2511	Sequence 1472511, A
C 19	35.6	4.2	1704	7 US-10-531-147-2511	Sequence 1472511, A
C 20	35.6	4.2	3593	12 US-60-742-219-2223	Sequence 2223, Ap
C 21	34.8	4.2	5157	7 US-10-461-673-7787	Sequence 7787, Ap

22	34.8	4.2	5174	10 US-11-266-748A-23500	Sequence 23500, A
23	34.2	4.1	449	8 US-11-239-610A-28539	Sequence 28539, A
24	34.2	4.1	5518	8 US-11-360-355-5183	Sequence 5183, Ap
25	34.2	4.1	1077	7 US-10-501-442-39	Sequence 39, Appl
26	34	4.1	4485	7 US-10-567-867-1977	Sequence 1977, Ap
27	34	4.1	4485	12 US-60-751-455-2720	Sequence 2720, Ap
28	33.8	4.1	5572	8 US-11-360-355-755	Sequence 755, App
29	33.8	4.1	118664	7 US-10-540-898-826	Sequence 826, App
30	33.6	4.1	7177	8 US-11-350-336-7	Sequence 7, Appl1
31	33.6	4.1	7364	8 US-11-350-336-5	Sequence 5, Appl1
32	33.6	4.1	7364	12 US-60-751-455-1947	Sequence 1947, Ap
33	33.6	4.1	7376	8 US-11-350-336-3	Sequence 3, Appl1
34	33.4	4.0	950	10 US-11-266-748A-357343	Sequence 357343, A
35	33.4	4.0	950	10 US-11-266-748A-440722	Sequence 440722, A
36	33.4	4.0	1062	7 US-10-560-224-40	Sequence 40, Appl
37	33.4	4.0	1062	7 US-10-560-414-40	Sequence 40, Appl
38	33.4	4.0	2110	6 US-10-533-519-1874	Sequence 1874, Ap
39	33.4	4.0	5834	7 US-10-567-867-2773	Sequence 2773, Ap
40	33.4	4.0	6132	7 US-10-567-867-2774	Sequence 2774, Ap
41	33.2	4.0	438	8 US-11-353-150-39324	Sequence 39324, A
42	33.2	4.0	221454	1 PCT-US05-18594-1	Sequence 1, Appl1
43	33	4.0	393	8 US-11-353-150-35925	Sequence 35925, A
44	33	4.0	576	8 US-11-329-388-9222	Sequence 9222, Ap
45	33	4.0	938	8 US-11-056-355B-1081	Sequence 1081, Ap

ALIGNMENTS

RESULT 1
US-10-018-470B-1/c
Sequence 1, Application US/10018470B
GENERAL INFORMATION:
APPLICANT: FRASER, Claire Marie
APPLICANT: HICKER, Erin Kathleen
APPLICANT: PETERSON, Jeremy D.
APPLICANT: TETTERLIN, Hevry
APPLICANT: VENTER, J. Craig
APPLICANT: MASTGANT, Vega
APPLICANT: GALBOTTI, Cesira
APPLICANT: MORA, Marirosa
APPLICANT: RATTI, Giulio
APPLICANT: SCARSELLI, Maria
APPLICANT: SCARIELLO, Vincenzo
APPLICANT: RAPPUOLI, Rino
APPLICANT: PIZZA, Mariagrazia
APPLICANT: GRANDI, Guido
TITLE OF INVENTION: NEISSERIA GENOMIC SEQUENCES AND METHODS OF THEIR USE
FILE REFERENCE: 223002100400
CURRENT APPLICATION NUMBER: US/10/018, 470B
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: PCT/US00/05928
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: US 60/132,068
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: GB 0004695.3
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/US99/23573
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 115
SOFTWARE: Seqwin99, version 1.0.4
SEQ ID NO 1
LENGTH: 2242716
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-018-470B-1
Query Match 99.8%; Score 826.4; DB 7; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 5.4e-24;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGCAAAACGACGTTATGAGCTTCGCCGACGAGACGAGGCGCACATTCGCGAT 60
|||||


```

Db      191 CAGCAGCTCGGGCACCAGAGCTGGGGCTTTGACCGTGGGATGAAGGGCTACTCAGGAC 250
QY      459 GCGGGCT 466
          |||||
Db      251 AGCAGGAT 258

RESULT 4
US-11-266-748A-12825/c
; Sequence 12825, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12825
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-12825

Query Match      4.3%; Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0,

QY      580 GGAATGCGGCTTGGCAGCTCAATCCCGCTTGTGCGCCCAAGAGCTGATTATGCCAAG 639
Db      247 GGAGGGGTGGCTACTACTCTGACCAAGCTCTCTGTGCAGCTGGCCCTGTGAGTGGCTTG 188

QY      640 TTTTCAGACACAAACAGGCGATTTGGGAGCCTGATCGAACAGACGCGCTCCTGAACGCG 699
Db      187 GGTCAAGGCCACACCTCTCCACTGAGCCCGGTGACAGAGCTTAGCGGCTCCTCTCAGCTTC 128

QY      700 AAACAGCAAGCGCGGATTTCCCGCCCAACATTTCAACACCGGCTGATCCGCGCTTGG 759
Db      127 TGGAGGACAGCGCGCGCTGCTGCTGCCACCACTGCTTCCAGACACTCTCTCCAGACTACCTAT 68
QY      760 ACCAAATCCAGCAGG 775
Db      67 CAGATCCACGACATG 52

RESULT 5
US-11-266-748A-64567
; Sequence 64567, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12825
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-12825

```

```

; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 64567
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-64567

Query Match          4.3%, Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

OY 580 GGAATGCCGGTTTGCACGCTCAATCCCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAG 639
Db 40 GGAGAGGGTGGCTACTACCTGACCAAGCCTCTCTGCCAGCCTGGCCCTGAGTGGCCTG 99
OY 640 TTTTACCAACCAAAACAGCGCATTTGGGCGACCTGATCCAAACGACCGCTCTTGAACCGC 699
Db 100 GGTGAGCGCCCAACACCTCTCCCACTGAGCCCGCTGAGAGCATGAGGCGCTCTTCAAGCCTG 159
OY 700 AAACAGCAAAAGCGCGATTTCCCGCCCAACATTTCAACACGCGCTGATCCGCGCCTTG 759
Db 160 TGGAGAGCAGCGCCCGCTGCGCTGCGACCACTGCTTCCAGACCTTCTCCGAGTAGCTTAT 219
OY 760 ACCAAATCAGCAGGG 775
Db 220 CAGATCCAGCAGTG 235

RESULT 6
US-11-266-748A-67399/c
; Sequence 67399, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Hartkin, Paul
; APPLICANT: Johnston, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662, 276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 64567
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-64567

Query Match          4.3%, Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO: 67399
;; LENGTH: 286
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
US-11-266-748A-67399

Query Match 4.3%; Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGGCAGCTCAATCCCGCTTGTGGCCCAAGAGCTGATTCGCAAG 639
DB 247 GGAAGGGGTGCTACTACTGACAGCCCTCTGCGAGCTGGCCCTGAGAGGCTTG 188
QY 640 TTTACGACCAAAACAGCGCATTTGGGAGCTGATGAAACAGCCGCTCTGAACGGC 699
DB 187 GGTACAGGCCCAACCTCTCCACTGAGCCCCGTGAGAGCTACGGCGCTCCCTCAGCCTC 128
QY 700 AAACGCAAAAGCGCGATTTCCCGCCCAACATTCAAACACCGCTGATCCGCGCTTG 759
DB 127 TGGGAGCAGCGCGCTGCTGCTCCACCCACTGCTTCACAGCACTCTCCGAGTAGCTAT 68
QY 760 ACCAAATCAGCAGG 775
DB 67 CAGGATCCGACAGTG 52

RESULT 7
US-11-266-748A-116165
; Sequence 116165, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 116165
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-116165

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGGCAGCTCAATCCCGCTTGTGGCCCAAGAGCTGATTCGCAAG 639

DB 146 GGAAGGGGTGCTACTACTACAGCCTCTCTGCGAGCTGGCCCTCTGAGTGGCTG 205
QY 640 TTTACGACCAAAACAGCGCATTTGGGAGCTGATGAAACAGCCGCTCTGAACGGC 699
DB 206 GGTACAGGCCCAACCTCTCCACTGAGCCCCGTGAGAGCTACGGCGCTCCCTCAGCCTC 265
QY 700 AAACGCAAAAGCGCGATTTCCCGCCCAACATTCAAACACCGCTGATCCGCGCTTG 759
DB 266 TGGGAGCAGCGCGCTGCTGCTCCACCACTGCTTCAGCACTCTCTCCGAGTAGCTAT 325
QY 760 ACCAAATCAGCAGG 775
DB 326 CAGGATCCGACAGTG 341

RESULT 8
US-11-266-748A-158329/c
; Sequence 158329, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 158329
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-158329

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGGCAGCTCAATCCCGCTTGTGGCCCAAGAGCTGATTCGCAAG 639
DB 855 GGAAGGGGTGCTACTACTGACAGCCCTCTGCGAGCTGGCCCTGAGAGGCTTG 796
QY 640 TTTACGACCAAAACAGCGCATTTGGGAGCTGATGAAACAGCCGCTCTGAACGGC 699
DB 795 GGTACAGGCCCAACCTCTCCACTGAGCCCCGTGAGAGCTACGGCGCTCCCTCAGCCTC 736
QY 700 AAACGCAAAAGCGCGATTTCCCGCCCAACATTCAAACACCGCTGATCCGCGCTTG 759
DB 735 TGGGAGCAGCGCGCTGCTGCTCCACCACTGCTTCAGCACTCTCTCCGAGTAGCTAT 676
QY 760 ACCAAATCAGCAGG 775
DB 675 CAGGATCCGACAGTG 660


```
RESULT 9
US-11-266-748A-286336
; Sequence 286336, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 286336
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-286336

Query Match
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGCGCCCAAGAGCTGATTATGCAAG 639
    |||
DB 146 GGAGAGGTGGTACTACTGACGACGCTCTGCGCAGCTGCGCTGAGTGAGCTG 205
    |||
QY 640 TTTCAGACCAAAACAGCGGATTTGGGAGCTGATGAAACAGACCGCTCTGAAACGC 639
    |||
DB 206 GGTACAGGCCACACCTCCCACTGAGCGCCGTCAGAGCTACGCGGCTCCCTCAGGCTC 265
    |||
QY 700 AAACAGCAAAAGCGGATTTCCCGGCAACATTTAAACAGCGGCTGATCGGCGCTTG 759
    |||
DB 266 TGGAGAGAGGCGCGCTGCTGCGACCCACTGCTTCAGACCTCTCCGAGTAGCTAT 325
    |||
QY 760 ACCAATATCAGCAGG 775
    |||
DB 326 CAGATCCAGCAGTG 341
    |||

RESULT 10
US-11-266-748A-337765/c
; Sequence 337765, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
```

```

; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 337765
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-337765

Query Match
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGCGCCCAAGAGCTGATTATGCAAG 639
    |||
DB 855 GGAGAGGTGGTACTACTGACGACGCTCTGCGCAGCTGCGCTGAGTGAGCTG 796
    |||
QY 640 TTTCAGACCAAAACAGCGGATTTGGGAGCTGATGAAACAGACCGCTCTGAAACGC 639
    |||
DB 795 GGTACAGGCCACACCTCCCACTGAGCGCCGTCAGAGCTACGCGGCTCCCTCAGGCTC 736
    |||
QY 700 AAACAGCAAAAGCGGATTTCCCGGCAACATTTAAACAGCGGCTGATCGGCGCTTG 759
    |||
DB 735 TGGAGAGAGGCGCGCTGCTGCGACCCACTGCTTCAGACCTCTCCGAGTAGCTAT 676
    |||
QY 760 ACCAATATCAGCAGG 775
    |||
DB 675 CAGATCCAGCAGTG 660
    |||

RESULT 11
US-11-266-748A-396766
; Sequence 396766, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 396766
; LENGTH: 1000
```

;; TYPE: DNA
;; ORGANISM: Homo Sapiens
US-11-266-748A-396766

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGCTTTCAGCTCAATCCCGCTTGTGGCCCAAGAGCTGATTATGCCAAG 639
DB 146 GGAAGAGGTGCTACTACTGCTGACAGCTCTCTGCGCAGCTGAGCTGAGTGGCTTG 205
QY 640 TTTCACGACCAAAACAGCGCATTTGGCGACGCTGATCGAACACAGCGCTCTGAACCGC 699
DB 206 GGTTCAGGCCACACCTCCCACTGAGCCCGTGCAGAGACTACGGGGCTCCCTCAGCCTC 265
QY 700 AAACGCAAAAGCGCGATTTCCCGCCCAACATTCAAACACCGCTGATCGCGGCTTG 759
DB 266 TGGAGACAGCGCGCTGCTGCTGCACCCACGCTCTTCAGACCTCTCCGAGTACCTAT 325
QY 760 ACCAAATCAGCAGG 775
DB 326 CAGATCCACAGATG 341

RESULT 12

US-11-266-748A-467812/c
; Sequence 467812, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patent version 3.3
; SEQ ID NO 467812
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-467812

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGCTTTCAGCTCAATCCCGCTTGTGGCCCAAGAGCTGATTATGCCAAG 639
DB 855 GGAAGAGGTGCTACTACTGCTGACAGCTCTCTGCGCAGCTGAGCTGAGTGGCTTG 796
QY 640 TTTCACGACCAAAACAGCGCATTTGGCGACGCTGATCGAACACAGCGCTCTGAACCGC 699
DB 795 GGTTCAGGCCACACCTCCCACTGAGCCCGTGCAGAGACTACGGGGCTCCCTCAGCCTTC 736

QY 700 AAACGCAAAAGCGCGATTTCCCGCCCAACATTCAAACACCGCTGATCGCGGCTTG 759
DB 735 TGGAGACAGCGCGCTGCTGCTGCACCCACGCTCTTCAGACCTCTCCGAGTACCTAT 676
QY 760 ACCAAATCAGCAGG 775
DB 675 CAGATCCACAGATG 660

RESULT 13

US-11-056-355B-89151
; Sequence 89151, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 89151
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2821)
; OTHER INFORMATION: Ceres Seq. ID no. 12715871
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2821)
; OTHER INFORMATION: Ortholog of Ceres SEQ ID NO 13590691
; OTHER INFORMATION: as cited in SEQ ID NO 66290
US-11-056-355B-89151

Query Match 4.3%; Score 36; DB 8; Length 2821;
Best Local Similarity 51.2%; Pred. No. 1.6;
Matches 84; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 639 GTTTCACGCAAAACAGCGATTTGGCGAGCTGATGAAACAGACGCGCTCTGAACCG 698
DB 1841 GTGGACACCAAAACAGACGACTGTGAGCGCTGATGAACATCTTCTTCCACAGGTCAA 1900
QY 699 CAACAGCAAAAGCGCGATTTCCCGCCCAACATTCAAACACGCGCTGATCGCGCTT 758
DB 1901 GTGAATCTAGATCGTTCTCGCGCAGCAAGCTAATGGAATTGTTAATGATCAGGAGTT 1960
QY 759 GACCAAAATCAGCAGGAAAGGAAACGCGCGCAAGGCGCG 802
DB 1961 GTATCTGCAATGCTGAAAMAGCTCAACCTGTGCAAAATGCGAG 2004

RESULT 14

US-11-056-355B-92907
; Sequence 92907, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 92907
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)-(2821)
OTHER INFORMATION: Ceres Seq. ID no. 12715871
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)-(2821)
OTHER INFORMATION: Ortholog of Ceres SEQ ID NO 13590691
OTHER INFORMATION: as cited in SEQ ID NO 66290
US-11-056-355B-92907

```

```

Query Match
Best Local Similarity 4.3%; Score 36; DB 8; Length 2821;
Matches 84; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

```

QY 639 GTTTCAGACCAAAACAGCCATTGGGACGCTGATGAAACAGACGCGCTCTGAACCG 698
DB 1841 GTGGCAGACCAAGACGACACTGTGAGCTGATGAAACATCTTTCACAGATCAA 1900
QY 699 CAACAGCAAGGCGGATTCCTCCGCAACACATTCACACGCGCTGATCCGCGCTT 758
DB 1901 GTGAATCTGATCGTTCTCGGACGCAAGCTAATGAAATGTTACTGATCAGGAGTT 1960
QY 759 GACCAAAATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 802
DB 1961 GTATCTGACAGATGCTGAAAAAGCTCAACCTGCTGCAAAATGCGAG 2004

```

```

RESULT 15
US-11-330-364-14690/C
Sequence 14690, Application US/11330364
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
APPLICANT: Andersen, Scott E.
APPLICANT: Dubois, Patrice E.
APPLICANT: Mahadeo, Debbie A.
APPLICANT: Masucci, James D.
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 16517,352 - 38-21(51721)C/US
CURRENT APPLICATION NUMBER: US/11/330,364
PRIOR FILING DATE: 2006-01-12
PRIOR APPLICATION NUMBER: US 09/696,664
PRIOR FILING DATE: 2000-10-25
PRIOR APPLICATION NUMBER: US 60/161,619
PRIOR FILING DATE: 1999-10-26
NUMBER OF SEQ ID NOS: 17472
SEQ ID NO 14690
LENGTH: 516
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3279-220-Q1-N6-A9
US-11-330-364-14690

```

```

Query Match
Best Local Similarity 4.3%; Score 35.6; DB 8; Length 516;
Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```

```

QY 279 CGGCGAAGGTAGAGAAAATTCCTGCGGAGAGGCTTGGCTGCAAGAAAGCTTTGACCC 338
DB 368 CGACGAGAGAGAGTGGTATCTCTGGGAAAGAGCGGACCTCAAGAGTGGCTGAGAGT 309
QY 339 GGAATACCGCTTATATGCTCGCTTGAAGAGATGTTATGACAGTCTGACCTCGCTTC 398
DB 308 CCGAAGTGAATGCTCGCTCGGAGCGTGAAGCTGAGTGGCGGACCTCCCTGCTCGT 249
QY 399 CGGCGTGGCGGATTAATCTGCGGCGGCGCTTTCGCTGTTGAAAGGAGAAACACTGGGGGAC 458
DB 248 CAGCACTGGGAGAGAGAGAGTGTCTGCTTTCAGTGAAGAGAGGAGGAGGAGGAGGAG 189
QY 459 GGGGGG 464

```

```

DB 188 GGCAGG 183
Search completed: April 7, 2006, 19:25:26
Job time : 3563.5 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:22:21 ; Search time 189 Seconds
(without alignments)
639.308 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MGNHYSLSAARRAHAD.....ISREPERKRRGRQFIVPRQ 275

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 244163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 244163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

A_Geneseq_21:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1446	100.0	275	2	AAV22156 N. mening
2	1288	89.1	279	2	AAK91312 N. gonorr
3	1281	88.6	279	6	ABP77390 N. gonorr
4	1007	69.6	280	2	AAK91315 N. gonorr
5	1002	69.3	280	2	AAW06580 Lipo-olig
6	996	68.9	280	6	ABP77398 N. gonorr
7	933	64.5	280	6	ABU06084 N. mening
8	364	25.2	255	8	ADU04721 M. catarr
9	251	17.6	50	2	AAW89328 Neisseria
10	235	16.0	50	2	AAW89329 Neisseria
11	130.5	9.0	521	3	AAK91300 Human ORF
12	130.5	9.0	622	4	AAW93427 Human pol
13	130.5	9.0	622	4	AAW88452 Human mem
14	130.5	9.0	622	5	ABP43477 Human sec
15	130.5	9.0	622	8	ADL31021 Human pro
16	130.5	9.0	622	9	ADY63267 Human c10
17	122.5	8.5	1128	4	ABW62795 Drosophila
18	121	8.4	738	8	AD018760 Human scf
19	118.5	8.2	423	5	ABW89453 Human pol
20	118.5	8.2	517	3	AAK81790 Human ORF
21	118.5	8.2	517	4	AAW93202 Human pol
22	118.5	8.2	517	4	AAW93206 Human pol
23	118.5	8.2	517	8	ADL30558 Human pro
24	118.5	8.2	517	8	ADL30556 Human pro

25	118.5	8.2	560	8	ADQ21291	Adg21291 Human sof
26	118.5	8.2	578	4	AAW81257	Aag81257 Human AFP
27	118.5	8.2	579	2	AAV30812	Aay30812 Human sec
28	118.5	8.2	594	6	ABU11815	Abu11815 Human MDD
29	118.5	8.2	594	6	ABU11514	Abu11514 Human MDD
30	118.5	8.2	595	3	AAK12121	Abk12121 Hydropob
31	118.5	8.2	536	5	ABP42016	Abp42016 Human ova
32	117.5	8.1	517	8	ABW81071	Abw81071 Tumour-as
33	113.5	7.8	527	8	ADH18902	Adh18902 Human cel
34	110	7.6	201	2	AAW20604	Aaw20604 H. pylori
35	110	7.6	739	4	ABW11417	Abw11417 Human sec
36	110	7.6	739	6	ABO00923	AbO00923 Polypept1
37	104	7.2	762	9	AEC08408	Aec08408 SCFV-B2-F
38	100.5	7.0	168	4	ABG22090	Abg22090 Novel hum
39	99.5	6.9	273	4	AAW60317	Aaw60317 Helicobac
40	99	6.8	234	8	ADP09098	Adp09098 Human pro
41	98.5	6.8	273	2	AAV32543	Aav32543 H. pylori
42	98.5	6.8	444	5	AAU76669	Aau76669 Helicobac
43	98	6.8	481	6	ABP60054	Abp60054 H. pylori
44	95.5	6.6	273	4	ABW60321	Abw60321 Helicobac
45	93	6.4	109	6	ABW54961	Abw54961 Mouse IgG

ALIGNMENTS

RESULT 1	AAV22156	standard; protein; 275 AA.
ID	AAV22156	
XX	AAV22156;	
AC	AAV22156;	
XX	AAV22156;	
DT	08-SEP-1999	(first entry)
XX	08-SEP-1999	
DE	N. meningitidis Beta-1,4-galactosyltransferase.	
XX	N. meningitidis Beta-1,4-galactosyltransferase.	
KW	Beta-1,4-galactosyltransferase; 1gtb; fusion protein; catalytic domain;	
KM	glycosyltransferase; accessory enzyme; nucleotide sugar formation;	
KW	saccharide donor; oligosaccharide synthesis;	
KM	carbohydrate structure development.	
XX	carbohydrate structure development.	
OS	Neisseria meningitidis.	
XX	Neisseria meningitidis.	
PN	W0931224-A2.	
XX	W0931224-A2.	
PD	24-JUN-1999.	
XX	24-JUN-1999.	
PF	15-DEC-1998;	98WC-CA001180.
XX	15-DEC-1998;	
PR	15-DEC-1997;	97US-0069443P.
XX	15-DEC-1997;	
PR	14-DEC-1998;	98US-00211691.
XX	14-DEC-1998;	
PA	(CANNA) NAT RES COUNCIL CANADA.	
XX	(CANNA) NAT RES COUNCIL CANADA.	
PI	Gilbert M, Young NM, Wakarchuk WW;	
XX	Gilbert M, Young NM, Wakarchuk WW;	
DR	WPI; 1999-395174/33.	
XX	WPI; 1999-395174/33.	
DR	N-PSDB; AAX84281.	
XX	N-PSDB; AAX84281.	
PT	A new glycosyltransferase fusion protein useful in the enzymatic	
XX	A new glycosyltransferase fusion protein useful in the enzymatic	
PT	synthesis of oligosaccharides.	
XX	synthesis of oligosaccharides.	
Example 2; Fig 2; 63pp; English.		
XX	Example 2; Fig 2; 63pp; English.	
CC	This sequence represents the Neisseria meningitidis Beta-1,4-	
XX	This sequence represents the Neisseria meningitidis Beta-1,4-	
CC	galactosyltransferase (also referred to as 1gtb). The invention relates	
XX	galactosyltransferase (also referred to as 1gtb). The invention relates	
CC	to a nucleic acid encoding a fusion protein that comprises a	
XX	to a nucleic acid encoding a fusion protein that comprises a	
CC	glycosyltransferase catalytic domain and a catalytic domain from an	
XX	glycosyltransferase catalytic domain and a catalytic domain from an	
CC	accessory enzyme that is involved in formation of a nucleotide sugar	
XX	accessory enzyme that is involved in formation of a nucleotide sugar	
CC	which is a saccharide donor for a glycosyltransferase reaction. The	
XX	which is a saccharide donor for a glycosyltransferase reaction. The	
CC	fusion protein is useful in the enzymatic synthesis of oligosaccharides.	
XX	fusion protein is useful in the enzymatic synthesis of oligosaccharides.	
CC	The fusion proteins are able to catalyse more than one reaction involved	
XX	The fusion proteins are able to catalyse more than one reaction involved	
CC	in the enzymatic synthesis. This is useful for the development of	
XX	in the enzymatic synthesis. This is useful for the development of	

CC therapeutic agents that have specific carbohydrate structures.
 CC Carbohydrates are involved in recognition elements on the surface of
 CC cells. The fusion protein can be used for the synthesis of both natural
 CC carbohydrates and synthetic derivatives with novel properties. The fusion
 CC polypeptide allows two glycosyltransferase reactions in a single vessel,
 CC provides improved yields of end products. Additionally, cleanup and
 CC disposal of extra solvents and by-products is reduced. The fusion protein
 CC can also use directly different donor analogues and various acceptors
 CC with a terminal galactose residue

SQ Sequence 275 AA;

Query Match 100.0%; Score 1446; DB 2; Length 275;
 Best Local Similarity 100.0%; Pred. No. 3,8e-144;
 Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNHVISLAAERRAHIAADTFRHGIPQFPDMLMSEERLEQMAELVPGISAHPIYLSG 60
 DB 1 MGNHVISLAAERRAHIAADTFRHGIPQFPDMLMSEERLEQMAELVPGISAHPIYLSG 60
 QY 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDVDVLLGEGEKEFLAEDAMLQERPDPTAFIVR 120
 DB 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDVDVLLGEGEKEFLAEDAMLQERPDPTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
 QY 181 VDLMMFSDPFREGMPVQCLNPALCAQELAHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
 DB 181 VDLMMFSDPFREGMPVQCLNPALCAQELAHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
 QY 241 PANTFKHRLIRALTIKISREERKRORRQRFIVPQ 275
 DB 241 PANTFKHRLIRALTIKISREERKRORRQRFIVPQ 275

RESULT 2

AAR91312 standard; protein; 279 AA.

AA91312;

DT 16-OCT-2003 (revised)
 DT 09-JUL-1996 (first entry)

DE N. gonorrhoeae glycosyltransferase IgtB.

XX Glycosyltransferase; lipo-oligosaccharide; lgt gene; LOS locus; vaccine.

OS Neisseria gonorrhoeae; strain F62.

XX WO9610086-A1.

PD 04-APR-1996.

PF 25-SEP-1995; 95WO-US012317.

PR 26-SEP-1994; 94US-00312387.

PA (UYRQ) UNIV ROCKEFELLER.

PI Gotachlich EC;

DR WPI; 1996-200924/20.

DR P-PSDB; AAT14061.

PT Nucleic acids encoding glycosyl transferase(s) - used in the diagnosis of
 PT infection with Neisseria and for the biosynthesis of oligosaccharide(s).

PS Claim 9; Fig 2c; 81pp; English.

XX 5 Glycosyltransferases (AAR91311-15) are products of the lgt locus

CC (AAT14061) of Neisseria gonorrhoeae strain F62. Glycosyltransferase IgtB
 CC (AAR91312) can be obtained by expression of the lgtB coding sequence in
 CC recombinant host cells. A method for adding Gal beta1-4 to GlcNAc or Glc
 CC comprises contacting a reaction mixture containing activated Gal to an
 CC acceptor moiety comprising a GlcNAc or Glc residue in the presence of
 CC IgtB. Oligosaccharides can be produced that, when attached to non-toxic
 CC lipids, are useful for Neisseria vaccine preparation. Blood group core
 CC oligosaccharides, and mimics of lacto-N-neotetraose, gangliosides and
 CC saccharide portions of globoglycolipids can also be produced using the
 CC enzymes. (Updated on 16-Oct-2003 to standardise OS field)

SQ Sequence 279 AA;

Query Match 89.1%; Score 1288; DB 2; Length 279;
 Best Local Similarity 90.0%; Pred. No. 2.1e-127;
 Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MGNHVISLAAERRAHIAADTFRHGIPQFPDMLMSEERLEQMAELVPGISAHPIYLSG 60
 DB 1 MGNHVISLAAERRAHIAADTFRHGIPQFPDMLMSEERLEQMAELVPGISAHPIYLSG 60
 QY 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDVDVLLGEGEKEFLAEDAMLQERPDPTAFIVR 120
 DB 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDVDVLLGEGEKEFLAEDAMLQERPDPTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
 QY 181 VDLMMFSDPFREGMPVQCLNPALCAQELAHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
 DB 181 VDLMMFSDPFREGMPVQCLNPALCAQELAHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
 QY 241 PANTFKHRLIRALTIKISREERKRORRQRFIVPQ 275
 DB 241 PANTFKHRLIRALTIKISREERKRORRQRFIVPQ 275

RESULT 3

ABP77390 standard; protein; 279 AA.

ABP77390;

DT 07-MAR-2003 (first entry)

DE N. gonorrhoeae amino acid sequence SEQ ID 1310.

XX Antibacterial; infection; vaccine; gene therapy.

OS Neisseria gonorrhoeae.

XX WO200279243-A2.

PD 10-OCT-2002.

PF 12-FEB-2002; 2002WO-IB002069.

PR 12-FEB-2001; 2001GB-00003424.

PA (CHIR-) CHIRON SPA.

PI Fontana MR, Pizza M, Masiagnani V, Monaci E;

DR WPI; 2003-058415/05.

DR N-PSDB; ABZ38360.

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.

PS Disclosure; Page 283; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.

CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX

SQ Sequence 279 AA;

Query Match 88.6%; Score 1281; DB 6; Length 279;
Best Local Similarity 90.0%; Pred. No. 1.1e-126;
Matches 251; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

QY 1 MONHVSLISAARRRAHIAITPGRHGIPPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
DB 1 MONHVSLISAARRRAHIAITPGSRGIPQFPDAMPSERLQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMSHAVLWKOALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
DB 61 VEKACFMSHAVLWKOALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
QY 121 LETFMHVLITSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
DB 121 LETFMHVLITSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
QY 181 VDIWMFSDFFDRBGMVCCQNPALCAQELHYAKFHONSLSLIEHDLNLRKQQRDS 240
DB 181 VDIWMFSDFFDRBGMVCCQNPALCAQELHYAKFHONSLSLIEHDLNLRKQQRDS 240
QY 241 PANTFKRLIRALTKISREKRRORREQ----FIVPQ 275
DB 241 PANTFKRLIRALTKISREKRRORREQ----FIVPQ 275

RESULT 4

AAR91315
ID AAR91315 standard; protein; 280 AA.

XX AAR91315;

XX 16-OCT-2003 (revised)
XX 09-JUL-1996 (first entry)

XX N. gonorrhoeae glycosyltransferase LgtE.

XX Glycosyltransferase; lipo-oligosaccharide; lgt gene; LOS locus; vaccine.

XX Neisseria gonorrhoeae; strain F62.

XX WO9610086-A1.

XX 04-APR-1996.

XX 25-SEP-1995; 95WO-US012317.

XX 26-SEP-1994; 94US-00312387.

XX (UYRO) UNIV ROCKEFELLER.

XX Gotschlich EC;

XX WPI; 1996-200924/20.

XX N-PSDB; AAT14061.

XX Nucleic acids encoding glycosyl transferase(s) - used in the diagnosis of
XX infection with Neisseria and for the biosynthesis of oligo-saccharide(s).

XX Claim 12; Fig 2f; 81pp; English.

XX 5 Glycosyltransferases (AAR91311-15) are products of the lgt locus
XX (AAT14061) of Neisseria gonorrhoeae strain F62. Glycosyltransferase LgtE
XX (AAR91315) can be obt'd. by expression of the lgtE coding sequence in

CC recombinant host cells. A method for adding Gal beta1-4 to GlcNAc or Glc
CC comprises contacting a reaction mixture containing activated Gal to an
CC acceptor moiety comprising a GlcNAc or Glc residue in the presence of
CC LgtE. Oligosaccharides can be produced that, when attached to non-toxic
CC lipids, are useful for Neisseria vaccine prep'n. Blood group core
CC oligosaccharides, and mimics of lacto-N-neotetraose, gangliosides and
CC saccharide portions of globoglycolipids can also be produced using the
CC enzymes. (Updated on 16-OCT-2003 to standardise OS field)
XX

SQ Sequence 280 AA;

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.2e-97;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVSLISAARRRAHIAITPGRHGIPPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
DB 1 MONHVSLISAARRRAHIAITPGSRGIPQFPDAMPSERLQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMSHAVLWKOALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
DB 61 VEKACFMSHAVLWKOALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
QY 121 LETFMHVLITSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
DB 121 LETFMHVLITSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
QY 181 VDIWMFSDFFDRBGMVCCQNPALCAQELHYAKFHONSLSLIEHDLNLRKQQRDS 240
DB 181 VDIWMFSDFFDRBGMVCCQNPALCAQELHYAKFHONSLSLIEHDLNLRKQQRDS 240
QY 241 PANTFKRLIRALTKISREKRRORREQ 269
DB 237 RSLKVMFDLKRALGKFKREKRRORREQ 265

RESULT 5

AAM06580
ID AAM06580 standard; protein; 280 AA.

XX AAM06580;

XX 17-OCT-2003 (revised)
XX 21-MAR-1997 (first entry)

XX lipo-oligosaccharide gene-encoded protein.

XX Polyglycosyltransferase; N-acetylglucosaminyl transferase;
XX N-acetylglucosaminyl transferase; lipo-oligosaccharide.

XX Neisseria gonorrhoeae; ATCC 33084.

XX WO9640971-A1.

XX 19-DEC-1996.

XX 03-JUN-1996; 96WO-US008323.

XX 07-JUN-1995; 95US-00478140.

XX (NEOS-) NEOSB TECHNOLOGIES INC.

XX Johnson KF, Roth S, Buczala SL;

XX WPI; 1997-052351/05.

XX N-PSDB; AAT49230.

XX Transfer of at least 2 saccharide units using polyglycosyltransferase -
XX isolated from N. gonorrhoeae, catalyses the addition of both GlcNAc and
XX GalNAc di-saccharide(s) units to a single galactose moiety.

XX Disclosure; Fig 2F-H; 38pp; English.

CC A lipooligosaccharide-encoding gene region (AA049230) of *Neisseria*
CC gonorrhoeae ATCC 33084 includes coding sequences for 5 proteins (AA006576
CC -80), one of which (AA006576) is a polylipoyltransferase that catalyses
CC the addition of GlcNAc and GalNAc disaccharides to a galactose moiety.
CC The function of the other 4 proteins is not stated in the specification.
CC (Updated on 17-OCT-2003 to standardise OS field)

CC XX Sequence 280 AA;

Query Match 69.3%; Score 1002; DB 2; Length 280;

Best Local Similarity 74.7%; Pred. No. 4e-97; Mismatches 45; Indels 4; Gaps 1;

Matches 201; Conservative 19; Mismatches 45; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASAEERAAHIAADTFGRHGIPQFPDAMPSEERLEQAAAEIVPGLSAHPYISG 60
DB 1 MGNHVISLSAASAEERAAHIAADTFGRHGIPQFPDAMPSEERLEQAAAEIVPGLSAHPYISG 60
QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDVLLGEGEKFLEADAMLOERPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDVLLGEGEKFLEADAMLOERPDPTAFIVR 120
QY 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGTAIYISRKAMRFFLDLRFALPPEGGLHP 180
DB 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGTAIYISRKAMRFFLDLRFALPPEGGLHP 180
QY 181 VDLMMFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFKHRLIRALTYISREERKRRORRQ 269
DB 241 PANTFKHRLIRALTYISREERKRRORRQ 269
QY 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265

RESULT 6
ABP77398
ID ABP77398 standard; protein; 280 AA.

XX ABP77398;

DT 07-MAR-2003 (first entry)

DE N. gonorrhoeae amino acid sequence SEQ ID 1336.

XX Antibacterial; infection; vaccine; gene therapy.

OS *Neisseria gonorrhoeae*.

XX WO200279243-A2.

PN 10-OCT-2002.

PF 12-FEB-2002; 2002WO-IB002069.

PR 12-FEB-2001; 2001GB-00003424.

PA (CHIR-) CHIRON SPA.

PI Fontana MR, Piazza M, Masignani V, Monaci E;

DR MPI; 2003-058415/05.

DR N-PSDB; AB238368.

PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
PT medicament for treating or preventing *N. gonorrhoeae* infection.

PS Disclosure; Page 284; 815pp; English.

CC The present invention relates to proteins from *Neisseria gonorrhoeae*.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*

CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records AB076736-AB081046 represent nucleic acid
CC molecules of the invention

CC XX Sequence 280 AA;

Query Match 68.9%; Score 996; DB 6; Length 280;

Best Local Similarity 74.3%; Pred. No. 1.7e-96; Mismatches 46; Indels 4; Gaps 1;

Matches 200; Conservative 19; Mismatches 46; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASAEERAAHIAADTFGRHGIPQFPDAMPSEERLEQAAAEIVPGLSAHPYISG 60
DB 1 MGNHVISLSAASAEERAAHIAADTFGRHGIPQFPDAMPSEERLEQAAAEIVPGLSAHPYISG 60
QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDVLLGEGEKFLEADAMLOERPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDVLLGEGEKFLEADAMLOERPDPTAFIVR 120
QY 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGTAIYISRKAMRFFLDLRFALPPEGGLHP 180
DB 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGTAIYISRKAMRFFLDLRFALPPEGGLHP 180
QY 181 VDLMMFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFKHRLIRALTYISREERKRRORRQ 269
DB 241 PANTFKHRLIRALTYISREERKRRORRQ 269
QY 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265

RESULT 7
AB06084
ID AB06084 standard; protein; 280 AA.

XX AB06084;

DT 23-OCT-2003 (revised)

DT 22-JAN-2003 (first entry)

DE N. meningitidis vaccine antigen #61.

XX Vaccine; antigen; meningococcal disease; pathogenic bacteria; meningitis.

XX *Neisseria meningitidis* serogroup B.

XX WO20027648-A2.

PN 03-OCT-2002.

PF 22-MAR-2002; 2002WO-GB001399.

PR 22-MAR-2001; 2001GB-00007219.

PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Robinson A, Goringe AR, Hudson MJ, Bracegirdle P, West DM;

PI Oliver KJ, Kroll JS, Langford PR;

DR MPI; 2003-018958/01.

DR N-PSDB; ABX09919.

PT Identifying an antigen for manufacturing a vaccine against meningococcal
PT infection, comprises contacting antibodies with polypeptides, detecting
PT polypeptide-antibody complexes, and identifying bound polypeptides as
PT antigens.

PS Claim 37; Page 279-280; 310pp; English.

CC The invention relates to identifying an antigen comprising: (a) obtaining
CC antibodies against a commensal bacteria, or an extract from a commensal
CC bacteria; (b) contacting the antibodies with polypeptides obtained from

CC an expression library of either a commensal or a pathogenic bacteria; (c)
CC determining whether the polypeptides bind to antibodies; and (d) (where a
CC polypeptide binds to an antibody) identifying that polypeptide as an
CC antigen. Also included are: (1) a method of preparing a vaccine
CC composition, comprising identifying an antigen with the above method, and
CC combining the antigen with a carrier; (2) a vaccine composition obtained
CC by the above methods; (3) an antigen identified by the above methods; (4)
CC a polypeptide encoded by all or a part of a nucleic acid sequence
CC comprising the *Neisseria lactamica* DNA sequences detailed in the
CC specification; (5) a vector comprising the nucleic acid molecule; (6) a
CC method of preparing a composition for vaccination against infection by
CC pathogenic bacteria, comprising: (a) obtaining a first antigen from a
CC commensal *Neisseria*; (b) comparing the amino acid sequence of the first
CC antigen with the amino acid sequence of the second antigen from a
CC pathogenic bacteria, or comparing the sequence of a nucleic acid which
CC codes for the first antigen with the sequence of the nucleic acid that
CC codes for the second antigen; and if the first antigen is homologous to
CC the second antigen or if the nucleic acid sequence for the first antigen
CC is homologous to that of the second antigen, and (c) preparing a
CC composition for vaccination against bacterial infection comprising the
CC first antigen; and (7) an antibody that binds to the polypeptide antigen.
CC The method is useful in screening commensal and pathogenic bacteria for
CC previously unidentified vaccine antigens by identifying polypeptide
CC antigens that bind to sera raised against commensal bacterial proteins.
CC The polypeptide is useful as a vaccine antigen which may be used in the
CC manufacture of a medicament for vaccination against meningococcal
CC infection (e.g. meningitis). The present sequence represents an antigenic
CC protein from the pathogenic bacteria *N. meningitidis*. (Updated on 23-OCT-
CC 2003 to standardise OS field)

SQ Sequence 280 AA;

Query Match 64.5%; Score 933; DB 6; Length 280;
Best Local Similarity 69.5%; Pred. No. 8.2e-90;
Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

QY 1 MONHVISLSAARRAHIAIDTFGRHGIIPQFPDAMPSERLBQAMALVGLSAHPYLSG 60
DB 1 MONHVISLSAARRAHIAIDTFGRHGIIPQFPDAMPSERLBQAMALVGLSAHPYLSG 60
QY 61 VERACFMSHAVLWKQALDEGLPYITVFEEDVLLGSEGEKFLADAMIQERPDPTAFIV 120
DB 61 VERACFMSHAVLWKQALDEGLPYITVFEEDVLLGSEGEKFLADAMIQERPDPTAFIV 120
QY 121 LRTFMHVLTPSPGVADYCGRAFPILSESHWGTAGYIISRKAMPFLDRPALPPEGILH 180
DB 121 LRTFMHVLTPSPGVADYCGRAFPILSESHWGTAGYIISRKAMPFLDRPALPPEGILH 180
QY 181 VDLMPSPDFPDRBEGMPCQINPALCAQELHYAKFDQNSALGSLIEHRLNKKOQRDS 240
DB 181 VDLMPSPDFPDRBEGMPCQINPALCAQELHYAKFDQNSALGSLIEHRLNKKOQRDS 240
QY 241 PANTFKHLIRALTKISREERKRRREQ 269
DB 241 PANTFKHLIRALTKISREERKRRREQ 269
DB 237 RSLKVMFDLKRALGKGRKKMERORQ 265

RESULT-8

ADL04721 .†.
ID ADL04721 standard; protein; 255 AA.

AC ADL04721;

XX 06-MAY-2004 (first entry)

XX M. catarrhalis protein #487.

XX Moraxella catarrhalis; infection.

XX Moraxella catarrhalis.

XX US6673910-B1.

PD 06-JAN-2004.
XX
XX 04-APR-2000; 2000US-00540236.
PF
XX 08-APR-1999; 99US-0128416P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Breton GL;
PI
XX
XX WPI; 2004-178127/17.
DR
XX N-PSDB; ADL02801.
DR
XX
XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
XX preparing a composition for diagnosing, preventing or treating infection
XX caused by Moraxella catarrhalis.
XX
XX disclosure; SEQ ID NO 2407; 429pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding an Moraxella
XX catarrhalis polypeptide. The nucleic acid is useful for preparing a
XX composition for diagnosing, preventing or treating infection caused by
XX Moraxella catarrhalis. The present sequence represents the amino acid
XX sequence of a M. catarrhalis protein.

SQ Sequence 255 AA;

Query Match 25.2%; Score 364; DB 8; Length 255;
Best Local Similarity 35.2%; Pred. No. 1.3e-29;
Matches 90; Conservative 49; Mismatches 103; Indels 14; Gaps 7;

QY 1 MONHVISLSAARRAHIAIDTFGRHGIIPQFPDAMPSERLBQAMALVGLSAHPYLSG 60
DB 3 IQNFVISTATKATRRHNCIFGKQIGAFEPDAVPTDISKXAQLSIPINNORLTDG 62
QY 61 VERACFMSHAVLWKQALDEGLPYITVFEEDVLLGSEGEKFLAE-DAMIQERPDPTAFIV 119
DB 63 -EKACGLSHVALMQQIDENLDYMAIFEDVYLGDSQGFLOELTITLQO----NAVDA 117
QY 120 RLRTFMHVLTPSPGVADYCGRAFPILSESHWGTAGYIISRKAMPFLDRPALPPEGILH 179
DB 118 KLET--VVERKIHKKAVTVLNHRQLCPLKTFHTGTAGYVISOQAKIILDYLSLDAFEPF 176
QY 180 PYDLMPSPDFPDRBEGMPCQINPALCAQELHYAKFDQNSALGSLIEHRLNKKOQRD 239
DB 177 PIDHVLFDALISK--MSVLQVNPAMVQ----AHVSEDDTFKSLIETORKQVNOHRRR 230
QY 240 SPANTFKHLIRALTK 255
DB 240 SPANTFKHLIRALTK 255
DB 231 TLAD-VGKRYYSIGK 245

RESULT 9

AAW89328
ID AAW89328 standard; peptide; 50 AA.

AC AAW89328;

XX 26-FEB-1999 (first entry)

XX *Neisseria meningitidis* IgG C-terminal peptide.XX *Neisseria meningitidis*; IgG; IgG; beta-1,4-galactosyltransferase;

XX glycosyltransferase; proteolytic enzyme.

XX *Neisseria meningitidis*.

XX WO9654331-A2.

XX 03-DEC-1998.

XX 26-MAY-1998; 98WO-IB000975.

PR 27-MAY-1997; 97US-0047751P.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Wakarchuk WW, Young NM;
 XX
 DR WPI; 1999-035177/03.
 XX
 PT Expressing high levels of glycosyltransferases - comprises use of either
 PT host cells deficient in proteolytic enzymes or modified
 PT glycosyltransferase genes deleted in a proteolytic recognition site.
 XX
 PS Example 1; Fig 8; 61pp; English.
 XX
 CC A method has been developed of expressing a glycosyltransferase in a host
 CC cell. The method comprises introducing into the host cell a nucleic acid
 CC encoding the glycosyltransferase and incubating the host cell under
 CC conditions appropriate for expression of the glycosyltransferase, where
 CC the host cell substantially lacks a protease that cleaves polypeptides
 CC between two consecutive positively charged amino acid residues. The
 CC glycosyltransferase can be used in vitro production of oligosaccharide
 CC structures which are potential therapeutic agents for use in the
 CC manipulation of cell-cell recognition events, particularly adhesion of
 CC bacteria and viruses to mammalian cells and leukocyte- endothelial cell
 CC interaction through selecting in inflammation. The method provides more
 CC readily recoverable active glycosyltransferases than prior art methods
 CC involving mammalian glycosyltransferases. The present sequence represents
 CC a C-terminal peptide from Neisseria meningitidis IgTB from the present
 CC invention
 CC
 SQ Sequence 50 AA;
 Query Match 17.6%; Score 255; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 4.6e-19;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 QY 226 EHDRLNKKQQRDPSANTFKRLRLALTLKIREREKRRQREPTVPPQ 275
 Db 1 EHDRLNKKQQRDPSANTFKRLRLALTLKISREREKRRQREPTVPPQ 50
 RESULT 10
 AAM89329
 ID AAM89329 standard; peptide; 50 AA.
 XX
 AC AAM89329;
 XX
 DT 26-FEB-1999 (first entry)
 XX
 DE Neisseria meningitidis IgTB C-terminal peptide.
 XX
 KW Neisseria meningitidis; IgTC; IgTB; beta-1,4-galactosyltransferase;
 KW glycosyltransferase; proteolytic enzyme.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9854331-A2.
 XX
 PD 03-DEC-1998.
 XX
 PF 26-MAY-1998; 98WO-1B000975.
 XX
 PR 27-MAY-1997; 97US-0047751P.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Wakarchuk WW, Young NM;
 XX
 DR WPI; 1999-035177/03.
 XX
 PT Expressing high levels of glycosyltransferases - comprises use of either
 PT host cells deficient in proteolytic enzymes or modified
 PT glycosyltransferase genes deleted in a proteolytic recognition site.

XX	Example 1; Fig 8; 61pp; English.
PS	
CC	A method has been developed of expressing a glycosyltransferase in a host
CC	cell. The method comprises introducing into the host cell a nucleic acid
CC	encoding the glycosyltransferase and incubating the host cell under
CC	conditions appropriate for expression of the glycosyltransferase, where
CC	the host cell substantially lacks a protease that cleaves polypeptides
CC	between two consecutive positively charged amino acid residues. The
CC	glycosyltransferase can be used in vitro production of oligosaccharide
CC	structures which are potential therapeutic agents for use in the
CC	manipulation of cell-cell recognition events, particularly adhesion of
CC	bacteria and viruses to mammalian cells and leukocyte-endothelial cell
CC	interaction through selecting in inflammation. The method provides more
CC	readily recoverable active glycosyltransferases than prior art methods
CC	involving mammalian glycosyltransferases. The present sequence represents
CC	a C-terminal peptide from <i>Neisseria meningitidis</i> 1gtb from the present
CC	invention
XX	
XX	Sequence 50 AA;
SO	
Query Match	16.0%; Score 231; DB 2; Length 50;
Best Local Similarity	94.0%; Pred. No. 1.6e-16;
Matches 47; Conservative	0; Mismatches 3; Indels 0; Gaps 0
Qy	226 EHDRLINRKOQRSDPANTFKRLIRALTKTSREREKRRORREQFIYDFQ 275
Db	1 EHDRLINRKOQRSDPANTFKRLIRALTKTSREREKRRORREQFIYDFQ 50
RESULT 11	
AAB41900	
ID	AAB41900 standard; protein; 521 AA.
XX	
AC	AAB41900;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human ORFX ORF1664 polypeptide sequence SEQ ID NO:3328.
XX	
KM	Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KM	vulnery; antiparasitic; antiparkinsonian; nootropic; neuroprotective;
KM	anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiact;
KM	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KM	hypotensive; dermatological; immunosuppressive; antiinflammatory;
KM	antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KM	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KM	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM	cholesterol ester storage; systemic lupus erythematosus; infection;
KM	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KM	bone damage; cartilage damage; antiinflammatory disease; coagulation;
KM	thrombosis; contraceptive.
XX	
OS	Homo sapiens.
XX	
PN	WO200058473-A2.
XX	
PD	05-OCT-2000.
XX	
PF	31-MAR-2000; 2000MO-USO08621.
XX	
PR	31-MAR-1999; 99US-0127607P.
PR	02-APR-1999; 99US-0127638P.
PR	05-APR-1999; 99US-0127728P.
PR	30-MAR-2000; 2000US-00540763.
XX	
PA	(CIRA-) CIRA GEN CORP.
XX	
PI	Shimkets RA, Leach M;
XX	
DR	WPI; 2000-602362/57.

DR N-PSDB; AAC761P9.

PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.

PS Claim 11; Page 2519-2520; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytosolic; hepatocellular; viralnary;
CC antiproliferic; antiparkinsonian; neurotropic; neuroprotective; osteoparthy;
CC anticonvulsant; antirheumatic; immunosuppressive; immunostimulant;
CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antineumatic; antihydrolyd; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease, to
CC enhance coagulation; to inhibit thrombolysis; and as a contraceptive

SQ Sequence 521 AA;

```
Query Match      9.0%; Score 130.5; DB 3; Length 521;
Best Local Similarity 9.1%; Score 130.5; DB 3; Length 521;
```

Best local similarity 25.1%; P-NO. 0.00019;
Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;

```

QY 5 VJSLASABRRHINIDTFGRHGIPRPFQFAL-----WPSREJQAMVEIVPGSAHNY--- 57
Dz 244 MINLRRDRRRRMLRALQAEIECRIVAEVNDGKAMNISOVEALGIOMLEPGY RDYFNGR 302
QY 58 -LSGV5KK6CM7SHAVLWKQALDEGLPYITVFEDDV-----LLSGEKEFLAED 104
Dz 303 PLTKSELGCFLSHNYIMKEVVDVRLQKSLVFEEDLLEFIFPKRLNMIMRDVE----- 355
QY 105 AMLQRRPBDPAFYRLFTFMFMHVLTSBGCVADYCSRAP---LLSESH-WGTAGYIIS 159
Dz 356 ---REGLDWDLTYVERKRMQVSH---PE-----KAPBRVNLVEADYSYWTLLAYVIS 401
QY 160 RKAMFFPLDRFALPPEGLHPVDLMFMFSDFPDRBEGMPCVQMLPALCAQELH 210
Dz 402 LQGARKL---NABPLSKRLPVDEBL-PVMFDGH---PVEEYKAHFEIRLNH 446

```

RESULT 12

ID AAM93427 standard; protein; 622 AA.

ID	standard; protein; 622 AA.
AA093427	

AC AAM93427;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3054.

KW Human; full length cDNA; cDNA synthesis; oligo-capping

OS Homo sapiens.

PN BP1130094-A2

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
PR

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T,

XX

DR N-PSDB; AAK94348.

PT 830 Primers useful for synthesizing full length cDNA clones and their use

2000

2. **XX**

The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

SQ Sequence 622 AA;

Query Match	9.08;	Score 130.5;	DB 4;	Length 622,
-------------	-------	--------------	-------	-------------

Best Local Similarity 25.1%; Pred. No. 0.00024;
Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12

```

QY 5 VYSLASAEARKAHADTFGRGICIPQPFQDAL----MPSERLEQAMALVYGLSANPR--- 57
DQ 345 MNINRRRDRORRMRRLRALQAQIEICRRLVEAVDGRKAMNTSQVEALGIQMLTNGY-RDPYHGR 4030
QY 58 -LSGYEKACFMNSHAYLMTKQALDEGLPYTYTFEDDV-----LLRSGEKKFLAED 1040
DQ 404 PLTKELGCGFLSHNYIMKEVAVDRLGQSLVYEDDLRREIIPPKRLMYMLMDVE----- 4560
QY 105 AMILOERFEDPDIATVRLFTMEAHVLTSPSGVADYCGRAFP-----LLESEH-WGTAGYIIS 1590
DQ 457 ---REGLDWDLTYVGRKQMYEH-----PE-----KAVFRVNVLYEADYSYTWLAVIS 5020
QY 160 RKAAMPFLLDRFAALPREGLHPVDLMTMSDFPDREGCMVQCQNLALCAQEH 210
DQ 503 LGQARKTL---AAEPLSTQLPEVDEFTL-PVMEDEKH--PVSEKKAHFLSLRNH 547

```

RESULT 13

ID AAB88452 standard; protein; 622 AA.

ID AAB88452 standard; protein; 622 AA.

AC AAB88452

DT 23-MAY-2001 (first entry)

DE Human membrane or secretory protein clone PSEC0241.

KW Human; secretory protein; membrane protein; vaccine; gene therapy;

XX

XX

100

XX

XX

Search completed: April 7, 2006, 15:25:52
Job time : 193 secs

```
RESULT 15
ADL31021
ID ADL31021 standard; protein; 622 AA.
XX
AC ADL31021;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein encoded by a full length cDNA clone SegID 3054.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method.
XX
OS Homo sapiens.
XX
PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; *93JP-00194486.
PR 11-JAN-2000; 2000JP-0018774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
DR N-PSDB; ADL31020.
XX
WIPI; 2004-204755/20.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Example 1; SEQ ID NO 3054; 1340bp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX
SQ Sequence 622 AA;
XX
Query Match 9.0%; Score 130.5; DB 8; Length 622;
Best Local Similarity 25.1%; Pred. No. 0.00024;
Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;
OY 1 5 VISLSAARRAHINDTFRHGIPQGFAL-----MPSERLEQAMAEIVPGSAHPY--- 57
DB 345 WILNRRRQDRRERMLRALDAQETBCRLVAVDGKANTISOEALGIQMLPGY-RDPYHGR 403
OY 58 -LSGVKACFMSHAVLWKOALDEGLPYITVFEDV-----LLGEGEKFLAED 104
DB 404 PLTKGELGFLSHYNWKEVNRGLQKSLVFEDDLRFELFFPKRLMNLMDYB----- 456
OY 105 ANLQERFDPDTAFIVRLTFMFWHLTSPGVADYCGRAF---LIESEH-WGTAGYIIS 159
DB 457 ---RBLGDWDLIYVGRKMQVEH---PR-----KAVPRVRLVEADYSYWTLAYVIS 502
OY 160 RKAMPFELRPAALPEGLHVPVLMWFSFPDREGMPVCOLNPAACAQELH 210
DB 503 LOGARKLL---AAEPLSKMLPVDEFL-PVMFDKH--PVSEYKAHPSLRNLH 547
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:26:11 ; Search time 40 Seconds
(without alignments)
661.490 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MONHVLSLSAARRAHAD.....ISRRERKRQRROFIVPQ 275

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1446	100.0	275	2	S70814	glycosyl transferase
2	1440	99.6	275	2	C81027	lacto-N-neotetraos
3	1324	91.6	279	2	A81971	lacto-N-neotetraos
4	1024.5	70.9	268	2	B81971	lacto-N-neotetraos
5	933	64.5	280	2	A81027	lacto-N-neotetraos
6	885	61.2	276	2	S70815	glycosyl transferase
7	454	31.4	266	2	S71025	lipopolysaccharide
8	410	28.4	302	1	A64077	lipopolysaccharide
9	297	20.5	282	1	F64091	lipopolysaccharide
10	202.5	14.0	268	2	AD3304	lipopolysaccharide
11	128.5	8.9	332	2	E71916	lipopolysaccharide
12	121	8.4	738	2	T00343	hypothetical prote
13	110.5	7.6	721	2	AH3417	lipopolysaccharide
14	106.5	7.4	267	2	B64175	hypothetical prote
15	102	7.1	492	2	T20368	hypothetical prote
16	99.5	6.9	273	1	B64623	lipopolysaccharide
17	99	6.8	611	2	I50715	A2 isoform of vacu
18	98.5	6.8	404	2	F71916	probable lipopolys
19	93.5	6.5	444	2	C64597	lipopolysaccharide
20	93	6.4	133	2	PC1155	ig heavy chain pre
21	91	6.3	284	2	G72578	probable ABC trans
22	91	6.3	1624	2	C71129	probable reverse g
23	90	6.2	738	2	A71462	probable glucan br
24	90	6.2	5069	2	T17464	rifamycin polyketi
25	89.5	6.2	379	2	F75144	sarcosine oxidase,
26	88.5	6.1	273	2	E71990	probable lipopolys
27	88.5	6.1	725	2	AF3336	3-methyl-2-oxobuta
28	88	6.1	284	1	B64620	lipopolysaccharide
29	88	6.1	742	2	D86160	hypothetical prote

30	85	5.9	256	2	T00097	hypothetical prote
31	85	5.9	292	2	D71894	probable lipopolys
32	85	5.9	1189	2	I39711	cell protein - Agr
33	84	5.8	785	2	T19741	hypothetical prote
34	83.5	5.8	653	2	AB1128	transcription anti
35	83.5	5.8	803	2	AD1282	leucyl-tRNA synthe
36	83.5	5.8	803	2	AH1653	leucyl-tRNA synthe
37	83	5.7	413	2	H88481	protein let-756 [1
38	83	5.7	9376	2	T14593	arylaminoacyl synth
39	82.5	5.7	384	2	A12962	cellulose synthesi
40	82.5	5.7	389	2	E98320	hypothetical prote
41	82	5.7	380	2	AC0511	probable carnitine
42	82	5.7	896	2	S53990	phycobilisome anch
43	82	5.7	963	2	S43344	phosphoenolpyruvat
44	82	5.7	2027	2	S60123	hypothetical prote
45	82	5.7	2056	2	G88564	protein R10B11.1 [

ALIGNMENTS

```
RESULT 1
S70814
glycosyl transferase B (BC 2.4.-.-) - Neisseria meningitidis
C/Species: Neisseria meningitidis
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #ext_change 09-Jul-2004
C/Accession: S70814
R/Vennings, M.P.; Hood, D.W.; Peak, I.R.A.; Virji, M.; Moxon, E.R.
Mol. Microbiol. 18, 729-740, 1995
A/Title: Molecular analysis of a locus for the biosynthesis and phase-variable expressi
A/Reference number: S70812; UID:96414473; PMID:8817494
A/Accession: S70814
A/Status: preliminary; nucleic acid sequence not shown
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-275 <JEN>
A/Cross-references: UNIPROT:Q51116; UNIPARC:UP1000016FCOF; EMBL:U25839; NID:9973183; PI
A/Genetics:
A/Genes: lgtB
C/Superfamily: lipopolysaccharide biosynthesis-associated protein
C/Keywords: glycosyltransferase

Query Match      100.0%; Score 1446; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 6e-122;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MONHVLSLSAARRAHADTFGRHGIFPQFDALMPSSRLDQAAELVPGLSAHPYLSG 60
      |||
DB      1 MONHVLSLSAARRAHADTFGRHGIFPQFDALMPSSRLDQAAELVPGLSAHPYLSG 60
      |||

QY      61 VRKACPMHVAIVMKQALDGLPYITVFPEDVILGGEKEFTLADNMLQRRFPDPAFYR 120
      |||
DB      61 VRKACPMHVAIVMKQALDGLPYITVFPEDVILGGEKEFTLADNMLQRRFPDPAFYR 120
      |||

QY      121 LETMFHVAITSPSGVADYCGRAFPILSESHMGCTAGIISRKAMRPFILDFEALPPREGLP 180
      |||
DB      121 LETMFHVAITSPSGVADYCGRAFPILSESHMGCTAGIISRKAMRPFILDFEALPPREGLP 180
      |||

QY      181 VDLMPFSDFFDEGMPVQALNPALCAQELHYAKFDONSALGSLIEHDLNRRKQQRDS 240
      |||
DB      181 VDLMPFSDFFDEGMPVQALNPALCAQELHYAKFDONSALGSLIEHDLNRRKQQRDS 240
      |||

QY      241 PANTFKHLIRALTITISRRERKRQRROFIVPQ 275
      |||
DB      241 PANTFKHLIRALTITISRRERKRQRROFIVPQ 275
      |||

RESULT 2
C81027
lacto-N-neotetraose biosynthesis glycosyl transferase lgtB NMB1928 [imported] - Neisseri
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #ext_change 09-Jul-2004
C/Accession: C81027
R/Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
```

Hickey, B.K.; Haft, B.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; PMID:20175755; PMID:10710307
A:Accession: CB1027
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <TET>
A:Cross-References: UNIPROT:Q51116, UNIPARC:UPI000012E5FD, GB:AE002541, GB:AE002098, NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1928
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 99.6%; Score 1440; DB 2; Length 275;
Best Local Similarity 99.6%; Pred. No. 2.1e-121;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MONHVLSASAERRAHIDTGRHGIPQFPDAMPSERLSQAMALVPGLSAHPYLSG 60
DB 1 MONHVLSASAERRAHIDTGRHGIPQFPDAMPSERLSQAMALVPGLSAHPYLSG 60

QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFPDPTAFIVR 120

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180

QY 181 VDLMPSPDFDEGMPVQCLNPAICAOELHYAKFHONSALSLIEHDLNRKQORRDS 240
DB 181 VDLMPSPDFDEGMPVQCLNPAICAOELHYAKFHONSALSLIEHDLNRKQORRDS 240

QY 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275
DB 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275

RESULT 3
A81971
lacto-N-neotetraose biosynthesis glycosyl transferase NMA0525 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: A81971
R:Patthill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; PMID:20222556; PMID:10761919
A:Accession: A81971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <PAR>
A:Cross-References: UNIPROT:P57033, UNIPARC:UPI000012E5FC, GB:AL162753, GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lgtB; NMA0525
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 91.6%; Score 1324; DB 2; Length 279;
Best Local Similarity 92.5%; Pred. No. 5.2e-111;
Matches 258; Conservative 2; Mismatches 15; Indels 4; Gaps 1;

QY 1 MONHVLSASAERRAHIDTGRHGIPQFPDAMPSERLSQAMALVPGLSAHPYLSG 60
DB 1 MONHVLSASAERRAHIDTGRHGIPQFPDAMPSERLSQAMALVPGLSAHPYLSG 60

QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFPDPTAFIVR 120

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180

QY 181 VDLMPSPDFDEGMPVQCLNPAICAOELHYAKFHONSALSLIEHDLNRKQORRDS 240
DB 181 VDLMPSPDFDEGMPVQCLNPAICAOELHYAKFHONSALSLIEHDLNRKQORRDS 240

QY 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275
DB 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275

RESULT 4
B81971
lacto-N-neotetraose biosynthesis glycosyl transferase NMA0527 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: B81971
R:Patthill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; PMID:20222556; PMID:10761919
A:Accession: B81971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <PAR>
A:Cross-References: UNIPROT:Q9JW65, UNIPARC:UPI00000C06, GB:AL162753, GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lgtA2; NMA0527
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 70.9%; Score 1024.5; DB 2; Length 268;
Best Local Similarity 73.6%; Pred. No. 3.3e-84;
Matches 209; Conservative 13; Mismatches 37; Indels 25; Gaps 4;

QY 1 MONHVLSASAERRAHIDTGRHGIPQFPDAMPSERLSQAMALVPGLSAHPYLSG 60
DB 1 MONHVLSASAERRAHIDTGRHGIPQFPDAMPSERLSQAMALVPGLSAHPYLSG 60

QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFPDPTAFIVR 120

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRFALPPEGGLHP 180

QY 181 VDLMPSPDFDEGMPVQCLNPAICAOELHYAKFHONSALSLIEHDLNRKQORRDS 240
DB 181 VDLMPSPDFDEGMPVQCLNPAICAOELHYAKFHONSALSLIEHDLNRKQORRDS 240

QY 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275
DB 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275

RESULT 5
A81027
lacto-N-neotetraose biosynthesis glycosyl transferase lgtE NMB1926 [imported] - *Neisseria*
C:Species: *Neisseria meningitidis*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81027
R:Belletijn, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, B.K.; Haft, B.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; PMID:20175755; PMID:10710307

A:Accession: A81027
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-280 <TEB>
 A:Cross-references: UNIPROT:Q51117; UNIPARC:UPI00000510B1; GB:A602541; GB:A602098; NID
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1926
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 64.5%; Score 933; DB 2; Length 280;
 Best Local Similarity 69.5%; Pred. No. 5.5e-76;
 Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 60
 1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 60
 61 VERKACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 120
 61 VERKACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 120
 121 LETMFHNVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 180
 121 LETMFHNVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 180
 181 VDMNMFTEFDEKGMFVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 181 VDMNMFTEFDEKGMFVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 241 PANTFKHLIRALTISRERERKQRRQ 269
 241 PANTFKHLIRALTISRERERKQRRQ 269
 237 RSLKVMFDLKRALGKGRKKRMRQRQ 265
 237 RSLKVMFDLKRALGKGRKKRMRQRQ 265

RESULT 6

S70815
 glycoyl transferase B (EC 2.4.4.-) - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
 C:Accession: S70815
 R:Jennings, M.P.; Hood, D.W.; Peak, I.R.A.; Virji, M.; Moxon, E.R.
 Mol. Microbiol. 18, 729-740, 1995
 A:Title: Molecular analysis of a locus for the biosynthesis and phase-variable expression
 A:Reference number: S70812; PMID:96414473; PMID:8617494
 A:Accession: S70815
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-276 <JEN>
 A:Cross-references: UNIPROT:Q51117; UNIPARC:UPI000016FC10; EMBL:U25839; NID:9773183; PID
 C:Genetics:
 A:Gene: jgtE
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein
 C:Keywords: glycosyltransferase

Query Match 61.2%; Score 885; DB 2; Length 276;
 Best Local Similarity 67.3%; Pred. No. 1.1e-71;
 Matches 181; Conservative 23; Mismatches 57; Indels 8; Gaps 2;

1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 60
 1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 60
 61 VERKACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 120
 61 VERKACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 120
 121 LETMFHNVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 180
 121 LETMFHNVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 180
 117 LETMFHNVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 176
 181 VDMNMFTEFDEKGMFVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 181 VDMNMFTEFDEKGMFVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240

Db 177 VDMNMFTEFDEKGMFVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 232
 QY 241 PANTFKHLIRALTISRERERKQRRQ 269
 Db 233 RSLKVMFDLKRALGKGRKKRMRQRQ 261

RESULT 7

S71025
 lipopolysaccharide biosynthesis protein lic2B - Haemophilus influenzae (strain RM7004)
 C:Species: Haemophilus influenzae
 C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S71025
 R:High, N.J.; Jennings, M.P.; Moxon, E.R.
 Mol. Microbiol. 20, 165-174, 1996
 A:Title: Tandem repeats of the tetramer 5'-CAAT-3' present in lic2A are required for ph
 A:Reference number: S71024; PMID:97014379; PMID:8861214
 A:Accession: S71025
 A:Molecule type: DNA
 A:Residues: 1-266 <HIG>
 A:Cross-references: UNIPROT:Q57394; UNIPARC:UPI000012B6CA; EMBL:Z54182; NID:G1150402; PI
 A:Experimental source: strain RM7004
 C:Genetics:
 A:Gene: lic2B
 C:Function:
 A:Description: required for the biosynthesis of a phase-variable lipopolysaccharide str
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 31.4%; Score 454; DB 2; Length 266;
 Best Local Similarity 39.7%; Pred. No. 4.2e-33;
 Matches 110; Conservative 42; Mismatches 107; Indels 18; Gaps 6;

QY 3 NNVISLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 62
 3 NNVISLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 62
 4 NNVISLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 62
 4 NNVISLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLSQMAELVPGLSAHPYLSG 62
 QY 63 KACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 122
 63 KACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 122
 Db 63 KACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 122
 63 KACFMSHAVLMKQALDGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIVR 122
 QY 123 TFMFHVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 182
 123 TFMFHVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 182
 Db 123 TFMFHVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 182
 123 TFMFHVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGILHP 182
 QY 183 LMEDFDFDRSGMPVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 183 LMEDFDFDRSGMPVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 Db 183 LMEDFDFDRSGMPVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 183 LMEDFDFDRSGMPVQVNPALCTQELHYAKFHNQNSALGSLIEHDLNLRKQQRDS 240
 QY 241 PANTFKHLIRALTISRERERKQRRQ 269
 241 PANTFKHLIRALTISRERERKQRRQ 269
 Db 233 ---TLKTVLISLAGKPKILKIKYKLFISKATVFR 266
 233 ---TLKTVLISLAGKPKILKIKYKLFISKATVFR 266

RESULT 8

A64077
 lipopolysaccharide biosynthesis protein lic2A - Haemophilus influenzae
 C:Species: Haemophilus influenzae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: A64077; S15282; S39576
 R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnness, E.F.; Kerlavage, J.
 ; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, C.
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; PMID:9530630; PMID:7542800
 A:Accession: A64077
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-302 <TIGR>
 A:Cross-references: UNIPROT:Q03974; UNIPARC:UPI000012E5B0; GB:U32736; GB:L42023; NID:915
 A:Experimental source: strain Rd K120
 R:Cope, L.D.; Yoge, R.; Mertsola, J.; Latimer, J.L.; Hanson, M.S.; McCracken Jr., G.H.;

Mol. Microbiol. 5, 1113-1124, 1991
A>Title: Molecular cloning of a gene involved in lipopolysaccharide biosynthesis and virulence in *Yersinia enterocolitica*
A/Reference number: S15282; MUID:92065807; PMID:1956289
A/Accession: S15282
A/Molecule type: DNA
A/Features: 1-28, 'N', 30-40, 45-105, 'U', 107-150, 'R', 152-302 <COP>
A/Cross-references: UNIPARC:UP1000016F7AB; EMBL:K56903; NID:948811; PIDD:CAA40221.1; PIDD:CAA40221.1; PIDD:CAA40221.1
A/Experimental source: strain 1642
A/Note: It is uncertain whether Met-1 or Met-11 is the initiator
R/Hit: N.J.; Deelman, M.B.; Moxon, E.R.
Mol. Microbiol. 9, 1275-1282, 1993
A>Title: The role of a repetitive DNA motif (5'-CAAT-3') in the variable expression of virulence in *Yersinia enterocolitica*
A/Reference number: S39576; MUID:95020659; PMID:7523834
A/Accession: S39576
A/Molecule type: DNA
A/Features: 1-28, 'N', 30-40, 49-105, 'U', 107-150, 'R', 152-255, 'P', 257-273, 'D', 275-302 <HIG>
A/Cross-references: UNIPARC:UP1000016F7B1; EMBL:K19441; NID:9305379; PIDD:AAA65534.1; PIDD:AAA65534.1; PIDD:AAA65534.1
A/Experimental source: strain RM7004
C/Genetics:
A/Function:
A/Description: involved in lipopolysaccharide biosynthesis and virulence expression
C/Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 28.4%; Score 410; DB 1; Length 302;
Best Local Similarity 32.2%; Pred. No. 4, 3e-29;
Matches 96; Conservative 59; Mismatches 113; Indels 30; Gaps 5;
1 MONVHSLASAEARRAHADTFGRHGIPOFPDAD-----MSESL 41
4 INIVISWENATERRRHGIPOFPDAD-----MSESL 63
42 EOMAAE---LVPLSHAPYLSGVEKAFMSHAYLMKQALDEGLPYTFVEDVLLGEGE 98
64 NOSINSNSILNHNIESRLITGEGKCLISHFLMKNKCNNEVEYKIFEDVLLGEGE 123
99 KPLAEDPATQERPDPTATIVLETFMAY-LTSPGVADYCGRAFPLESEHMGAGYI 157
124 VFLNOMEWKTRDFPDITIRLETFLQPVKEKQKRIPPFSRNDDIKSTWGTAGYI 183
158 ISRKARFPLDRPALPPEGLHPVDLMFSDPFRGMPVQOLPALCAQELHYAKFHQ 217
184 ISGAKAYIVELYKIPSEIYAVDELINFKLVDVNYIYQNLPAICIQEL---QANQ 240
QY 218 NSLGLIETHDLNLRKQQRSPANTPKRLIRALTKISREKRRQRRQPIVFO 275
DB 241 KSVLTGSLERKQKSKIRK---KTLKQRLTRIKENIRALNKKMKQKIKEMQ 294

RESULT 9
F64091
Lipopolysaccharide biosynthesis protein 11c2B homolog - Haemophilus influenzae (strain R
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C/Accession: F64091
R/Features: 1-28, 'N', 30-40, 45-105, 'U', 107-150, 'R', 152-302 <COP>
A/Cross-references: UNIPARC:UP10000139C94; EMBL:K56903; NID:948811; PIDD:CAA40221.1; PIDD:CAA40221.1; PIDD:CAA40221.1
A/Experimental source: strain 1642
A/Note: It is uncertain whether Met-1 or Met-11 is the initiator
R/Hit: N.J.; Deelman, M.B.; Moxon, E.R.
Mol. Microbiol. 9, 1275-1282, 1993
A>Title: The role of a repetitive DNA motif (5'-CAAT-3') in the variable expression of virulence in *Yersinia enterocolitica*
A/Reference number: S39576; MUID:95020659; PMID:7523834
A/Accession: S39576
A/Molecule type: DNA
A/Features: 1-28, 'N', 30-40, 49-105, 'U', 107-150, 'R', 152-255, 'P', 257-273, 'D', 275-302 <HIG>
A/Cross-references: UNIPARC:UP10000139C94; EMBL:K19441; NID:9305379; PIDD:AAA65534.1; PIDD:AAA65534.1; PIDD:AAA65534.1
A/Experimental source: strain RM7004
C/Genetics:
A/Function:
A/Description: involved in lipopolysaccharide biosynthesis and virulence expression
C/Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 20.5%; Score 297; DB 2; Length 282;
Best Local Similarity 31.0%; Pred. No. 5, 3e-19;

Matches 86; Conservative 44; Mismatches 115; Indels 32; Gaps 11;
QY 3 NHVHSLASAEARRAHADTFGRHGIPOFPDAD-----MSESL 41
DB 34 NVHSLTEQRRHRIIEFGKONIPEFPALTP-DIEETAKFNITLRSKALSD 92
QY 61 VEKAFMSHAYLMKQALDEGLPYTFVEDVLLGEGEKPLAEDPATQERPDPTATIV 120
DB 93 GEICGASHIVLMQALDENNNYINIFEDDITLGENNAKELLEID-----YISDDIHYLK 146
QY 121 LEFMHAYLTPSPGVADYCGR-APELSEHMGAGYIISRKARFPLDRPALPPEGLH 179
DB 147 LEANGKFFKQPKSVK-CDNNTVETVYKQ-SGAGYVTKGAKYILVYKNEPLD--V 201
QY 180 PVDLMFSDPFRGMPVQOLPALCAQELHYAKFHQNSALGSLIE-HDPLNKKQRR 238
DB 202 AVDSIVFEDFLHFKDYKIVQLSPGICVQDF---VLHPDNPPRESSIOEGRDV--HGNQK 256
QY 239 DSPANTFKHRLIRALTKISREKRRQRRQPIVFO 275
DB 257 SILEKIKNBERGKVKIKMGKQ-----VPEK 282

RESULT 10
AD3304
Lipopolysaccharide biosynthesis protein 11c2B (imported) - Brucella melitensis (strain
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AD3304
R/Features: 1-268 <KUN>
A/Cross-references: UNIPARC:UP10000057C9D; EMBL:K56903; NID:948811; PIDD:CAA40221.1; PIDD:CAA40221.1; PIDD:CAA40221.1
A/Experimental source: strain 164
C/Genetics:
A/Function:
A/Description: involved in lipopolysaccharide biosynthesis and virulence expression
C/Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 14.0%; Score 202.5; DB 2; Length 268;
Best Local Similarity 28.4%; Pred. No. 1, 4e-10;
Matches 74; Conservative 34; Mismatches 88; Indels 65; Gaps 11;
QY 1 MONVHSLASAEARRAHADTFGRHGIPOFPDAD-----MSESL 41
DB 9 MCVYILINDKSRDRLEFMAOFRLOAOF-----BRVAVNGRAMSPLELASFTQI 59
QY 58 -----LSGVEKAFMSH-AVLMKQALDEGLPYTFVEDVLLGEGEKPLAEDPATQ 109
DB 60 SKEMWAPLSPALIGGFLSHKCKLEKIAGED-AVAVFEDDIRLSQSSRFPLASDHVI-- 116
QY 110 RDPDPTATIVLETFMAYLTPSPGVADYCGRAFPLESEHMGAGYIISRKARFPL 167
DB 117 ---PKQADIVKIDAGHEVLIS-NPKYKGPYSIRLRSRHLQDTGGYVSRDARKLPL 172
QY 168 -----DRFALPPEGLHPVDLMFSDPFRGMPVQOLPALCAQELHYAKFHQNS 219
DB 173 MEKVASPVVHFLPDDPDG-----PFDNF-----EIVQISDALCROS-----GMS 212
QY 220 ALGSLIETHDLNLRKQQRSDS 240
DB 213 TIGQ-----NREKQKPS 225

RESULT 11
E71916
Probable lipopolysaccharide biosynthesis protein - Helicobacter pylori (strain J99)
C/Species: Helicobacter pylori

RESULT 12
T00343
hypothetical protein KIAA0584 - human (fragment)
C|Species: Homo sapiens (man)
C|Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C|Accession: T00343
R|Nasease, T.; Ihikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A|Title: Prediction of the coding sequences of unidentified human genes. IX. The complet
A|Reference number: Z14086; MUID:58290545; PMID:9628581
A|Accession: T00343
A|Status: preliminary; translated from GB/EMBL/DBDJ
A|Molecule type: mRNA
A|Residues: 1-738 <NAG>
A|Cross-references: UNIPROT:O60327; UNIPARC:UPI000006D2E4; EMBL:AB011156; NID:G3043691;
A|Experimental source: brain
C|Genetics:
A|Note: KIAA0584

Query Match

Best Local Similarity 7.6%; Score 110.5; DB 2; Length 721;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

Query 5 VTSLSAASERRAHADTFGRHGILPPOF-----PDALMPSERLQ 43
DB 25 VINNASOPAAVYKVVASIEAVYGGFOLHRIDAVNGHATGRTORIGIDARFPAAG----- 78
Query 44 AMAELVPGSLAHPIYLSGVKACFMASHAVLWKQALDEGLPYITVPEDVDVLLGEGBEKPLAE 103
DB 79 --RMDLPG-----EYGYTRSHSKALSESTLSDSPGLGILLEDVDVVTETTSARIIHD 126
Query 104 DAWLQERFDDPTAFIVL-----ETMFHVLVTPSGVADYCGARFAPLLESEHWGAGYIIS 159
DB 127 IIRSLPDPD-----VVLVLVNRSPFLPWSLLETAG--DRIGRA---IHGQGSAAVLYS 176
Query 160 RKAMRFPLDRPAL 173
DB 177 REGARKLLSALSTM 190

RESULT 14

EC64175

hypothetical protein H11697 (Jsg locus) - Haemophilus influenzae

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C/Accession: E64175; S27580

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirtness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J. D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter, A./Title: Whole-genome mapping and assembly of Haemophilus influenzae Rd.

A/Reference number: A64000; MUID:95350630; PMID:7542800

A/Accession: E64175

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-267 <TIGR>

Query Match

Best Local Similarity 7.6%; Score 110.5; DB 2; Length 721;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

Query 5 VTSLSAASERRAHADTFGRHGILPPOF-----PDALMPSERLQ 43
DB 25 VINNASOPAAVYKVVASIEAVYGGFOLHRIDAVNGHATGRTORIGIDARFPAAG----- 78
Query 44 AMAELVPGSLAHPIYLSGVKACFMASHAVLWKQALDEGLPYITVPEDVDVLLGEGBEKPLAE 103
DB 79 --RMDLPG-----EYGYTRSHSKALSESTLSDSPGLGILLEDVDVVTETTSARIIHD 126
Query 104 DAWLQERFDDPTAFIVL-----ETMFHVLVTPSGVADYCGARFAPLLESEHWGAGYIIS 159
DB 127 IIRSLPDPD-----VVLVLVNRSPFLPWSLLETAG--DRIGRA---IHGQGSAAVLYS 176
Query 160 RKAMRFPLDRPAL 173
DB 177 REGARKLLSALSTM 190

RESULT 13

AH3417

Ipsa protein [imported] - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C/Accession: AH3417

R/DeVecchio, V.G.; Kapratral, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, B.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leteski Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitense

A/Reference number: AD3252; PMID:11756688

A/Accession: AH3417

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-721 <KUR>

A/Cross-references: UNIPROT:Q8YG36; UNIPARC:UP10000058017; GB:AE008917; PIDD:AA152507.1, A/Experimental source: strain 16M

C/Genetics:

A/Gene: BMH1326

A/Map position: 1

Query Match

Best Local Similarity 7.6%; Score 110.5; DB 2; Length 721;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

Query 5 VTSLSAASERRAHADTFGRHGILPPOF-----PDALMPSERLQ 43
DB 25 VINNASOPAAVYKVVASIEAVYGGFOLHRIDAVNGHATGRTORIGIDARFPAAG----- 78
Query 44 AMAELVPGSLAHPIYLSGVKACFMASHAVLWKQALDEGLPYITVPEDVDVLLGEGBEKPLAE 103
DB 79 --RMDLPG-----EYGYTRSHSKALSESTLSDSPGLGILLEDVDVVTETTSARIIHD 126
Query 104 DAWLQERFDDPTAFIVL-----ETMFHVLVTPSGVADYCGARFAPLLESEHWGAGYIIS 159
DB 127 IIRSLPDPD-----VVLVLVNRSPFLPWSLLETAG--DRIGRA---IHGQGSAAVLYS 176
Query 160 RKAMRFPLDRPAL 173
DB 177 REGARKLLSALSTM 190

RESULT 13

AH3417

Ipsa protein [imported] - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C/Accession: AH3417

R/DeVecchio, V.G.; Kapratral, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, B.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leteski Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitense

A/Reference number: AD3252; PMID:11756688

A/Accession: AH3417

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-721 <KUR>

A/Cross-references: UNIPROT:Q8YG36; UNIPARC:UP10000058017; GB:AE008917; PIDD:AA152507.1, A/Experimental source: strain 16M

C/Genetics:

A/Gene: BMH1326

A/Map position: 1

Query Match

Best Local Similarity 7.6%; Score 110.5; DB 2; Length 721;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

Query 5 VTSLSAASERRAHADTFGRHGILPPOF-----PDALMPSERLQ 43
DB 25 VINNASOPAAVYKVVASIEAVYGGFOLHRIDAVNGHATGRTORIGIDARFPAAG----- 78
Query 44 AMAELVPGSLAHPIYLSGVKACFMASHAVLWKQALDEGLPYITVPEDVDVLLGEGBEKPLAE 103
DB 79 --RMDLPG-----EYGYTRSHSKALSESTLSDSPGLGILLEDVDVVTETTSARIIHD 126
Query 104 DAWLQERFDDPTAFIVL-----ETMFHVLVTPSGVADYCGARFAPLLESEHWGAGYIIS 159
DB 127 IIRSLPDPD-----VVLVLVNRSPFLPWSLLETAG--DRIGRA---IHGQGSAAVLYS 176
Query 160 RKAMRFPLDRPAL 173
DB 177 REGARKLLSALSTM 190

RESULT 13

AH3417

Ipsa protein [imported] - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C/Accession: AH3417

R/DeVecchio, V.G.; Kapratral, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, B.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leteski Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitense

A/Reference number: AD3252; PMID:11756688

A/Accession: AH3417

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-721 <KUR>

A/Cross-references: UNIPROT:Q8YG36; UNIPARC:UP10000058017; GB:AE008917; PIDD:AA152507.1, A/Experimental source: strain 16M

C/Genetics:

A/Gene: BMH1326

A/Map position: 1

Query Match

Best Local Similarity 7.6%; Score 110.5; DB 2; Length 721;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

Query 5 VTSLSAASERRAHADTFGRHGILPPOF-----PDALMPSERLQ 43
DB 25 VINNASOPAAVYKVVASIEAVYGGFOLHRIDAVNGHATGRTORIGIDARFPAAG----- 78
Query 44 AMAELVPGSLAHPIYLSGVKACFMASHAVLWKQALDEGLPYITVPEDVDVLLGEGBEKPLAE 103
DB 79 --RMDLPG-----EYGYTRSHSKALSESTLSDSPGLGILLEDVDVVTETTSARIIHD 126
Query 104 DAWLQERFDDPTAFIVL-----ETMFHVLVTPSGVADYCGARFAPLLESEHWGAGYIIS 159
DB 127 IIRSLPDPD-----VVLVLVNRSPFLPWSLLETAG--DRIGRA---IHGQGSAAVLYS 176
Query 160 RKAMRFPLDRPAL 173
DB 177 REGARKLLSALSTM 190

RESULT 13

AH3417

Ipsa protein [imported] - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C/Accession: AH3417

R/DeVecchio, V.G.; Kapratral, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, B.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leteski Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitense

A/Reference number: AD3252; PMID:11756688

A/Accession: AH3417

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-721 <KUR>

A/Cross-references: UNIPROT:Q8YG36; UNIPARC:UP10000058017; GB:AE008917; PIDD:AA152507.1, A/Experimental source: strain 16M

C/Genetics:

A/Gene: BMH1326

A/Map position: 1

Query Match

Best Local Similarity 7.6%; Score 110.5; DB 2; Length 721;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

Query 5 VTSLSAASERRAHADTFGRHGILPPOF-----PDALMPSERLQ 43
DB 25 VINNASOPAAVYKVVASIEAVYGGFOLHRIDAVNGHATGRTORIGIDARFPAAG----- 78
Query 44 AMAELVPGSLAHPIYLSGVKACFMASHAVLWKQALDEGLPYITVPEDVDVLLGEGBEKPLAE 103
DB 79 --RMDLPG-----EYGYTRSHSKALSESTLSDSPGLGIL

A:Experimental source: strain Rd KM20
 R:McLaughlin, R.; Abu Kwak, Y.; Spinoia, S.; Apicella, M.
 Submitted to the EMBL Data Library, June 1992
 A:Description: Characterization and sequence of the lsg locus from Haemophilus influenza
 A:Reference number: S27577
 A:Accession: S27580
 A:Molecule type: DNA
 A:Residues: 11-26, 'K', 28-39, 'I', 41-55, 'S', 57-111, 'L', 113-136, 'N', 138-139, 'N', 141-143, 'T'
 A:Cross-references: UNIPARC:UPI000016F78F; EMBL:M94855; NID:g148931; PIDN:AAA24981.1; PI
 A:Experimental source: strain A2

Query Match 7.4%; Score 106.5; DB 2; Length 267;
 Best Local Similarity 23.4%; Pred. No. 0.057; Indels 43; Gaps 14;
 Matches 60; Conservative 46; Mismatches 107;

QY 1 MGNHYISLASAERRAHADTF--GRHGIPQFPDAMPSERLQAMAE--VPGISAP 56
 DB 12 LKKYILSLDKDIOQR---ELFFSQKNTEDFQVFSAINTMQKWDLAIFNIEQKX 66
 QY 57 YLSGVEK--ACFMSHAVIMKQALDEG---LPYITVFEDDYLIGGEKFLAEDAMLOE 109
 DB 67 YFRNVTKGRICTLSHYOKIVEDNDIABDSYALVCEDDALFHSDFQONLT--ALLSE 124
 QY 110 RPDPTAFI-----VRLETWFMHVLTPSGVADYCGRAFPILSEHWGTAG 155
 DB 125 KLBAGIILLGSKINDRNDPDLINYPITF-SFLCKKTGDVNY--AFP-YKSYFAGTVG 179
 QY 156 YIISRRAMPFLDRFAPLPBGLHPVDLMMFSDPFDRBGMFVQOL---NPALCAQELHY 211
 DB 180 YLIKSGAARFICQISQNKRFWLAD-DFLIFBQNFIRNKVRPLVIEHPVLIS-NLES 237
 QY 212 AKFHQNSALGSLIEH 227
 DB 238 VRGSLSNMLKLMKY 253

RESULT 15
 T20368
 hypothetical protein D2045_9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T20368
 R:Lloyd, C.
 submitted to the EMBL Data Library, August 1994
 A:Reference number: Z19262
 A:Accession: T20368
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-492 <WIL>
 A:Cross-references: UNIPROT:Q18989; UNIPARC:UPI0000079D6A; EMBL:Z35639; PIDN:CAA84699.1;
 A:Experimental source: clone D2045
 C:Genetics:
 A:Gene: CESP:D2045.9
 A:Map position: 3
 A:Introns: 87/2; 140/3; 291/2; 367/2

Query Match 7.1%; Score 102; DB 2; Length 492;
 Best Local Similarity 21.2%; Pred. No. 0.31; Indels 52; Gaps 9;
 Matches 47; Conservative 35; Mismatches 88;

QY 4 HVISLASAERRAHADTFGRHGIPQFPDAMPSERLQAMAE-----VPGISA 54
 DB 293 YLVNLKRQERLNMOKIFDLISIEYSLER-TDQKIDELPELKNYQILLEGIDPISK 351
 QY 55 HPYLSGVEKACFMSHAVIMKQALDEGLPYITVFEDDYLIGGEKFLAEDAMLOERFDPD 114
 DB 352 RPKMGK-BIGCFLSHYRWQDVVOHNYEKYIVFEDDL-----RPSHD 392
 QY 115 TAPYRLETMFHVLTPSGVADYCGRA-----PILSEHWGTAGYIIS 159
 DB 393 GLTRIRREVLDLDSHKFWDLI-YGRKKQSENEELMISQRRHLSVFIYSYV-TLGTMLS 450
 QY 160 RKAMRFLDRFAPLPBGLHPVD--LMMFSDPFDRBGMFVQ 198

DB 451 LMGARKIL--RPNPLKQWVPVDEYLPIMFNKHPNKVSLKIC 489
 Search completed: April 7, 2006, 15:30:35
 Job time : 42 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:22:56 ; Search time 230 Seconds

(without alignments)
843.566 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHVISLMSAERRAHAD.....ISRREKRRGRQRPVFPQ 275

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	99.6	275	1 LGTB_NEIME	O5116 neisseria m
2	1427	98.7	275	2 O8KR92_NEIME	O8KR92 neisseria m
3	1399	96.7	279	2 O7WUX5_NEISU	O7WUX5 neisseria s
4	1392	96.3	275	2 O8L2V3_NEIME	O8L2V3 neisseria m
5	1391	96.2	275	2 O8L2V5_NEIME	O8L2V5 neisseria m
6	1386	95.9	275	2 O8L2V7_NEIME	O8L2V7 neisseria m
7	1361	94.1	275	2 O9EVD4_NEISU	O9EVD4 neisseria s
8	1360	94.1	279	2 O8K1A8_NEIME	O8K1A8 neisseria m
9	1353	93.6	279	2 O7WUX2_NEIPO	O7WUX2 neisseria p
10	1333	92.2	279	2 O8L2V9_NEIME	O8L2V9 neisseria m
11	1328	91.8	275	2 O8K1G8_NEIME	O8K1G8 neisseria m
12	1327	91.8	275	2 O8L2V2_NEIME	O8L2V2 neisseria m
13	1325	91.6	279	2 O8L2V6_NEIME	O8L2V6 neisseria m
14	1324	91.6	279	1 LGTB_NEIMA	B57033 neisseria m
15	1320	91.3	275	2 O93E5_NEIME	O93E5 neisseria m
16	1288	89.1	279	1 LGTB_NEIGO	O50947 neisseria g
17	1281	88.6	279	2 O5F4Y6_NEIG1	O5F4Y6 neisseria g
18	1273	88.0	279	2 O7WUX9_NEILA	O7WUX9 neisseria l
19	1250	86.4	279	2 O93PR9_NEIGO	O93PR9 neisseria i
20	1187	82.1	248	2 O9RGN1_NEIGO	O9RGN1 neisseria g
21	1122.5	77.6	268	2 O8KR91_NEIME	O8KR91 neisseria m
22	1055	73.0	280	2 O93PR7_NEIGO	O93PR7 neisseria g
23	1044.5	72.2	268	2 O8K115_NEIME	O8K115 neisseria m
24	1033.5	71.5	268	2 O8K1G7_SWINE	O8K1G7 neisseria m
25	1024.5	70.9	268	2 O7BBB1_NEIME	O7BBB1 neisseria m
26	1020	70.5	268	2 O9JW65_NEIMA	O9JW65 neisseria m
27	1017.5	70.4	280	2 O8L2V8_NEIME	O8L2V8 neisseria m
28	1017.5	70.4	268	2 O7WUX7_NEILA	O7WUX7 neisseria l
29	1017.5	70.4	268	2 O8L2U9_NEIME	O8L2U9 neisseria m
30	1013.5	70.1	268	2 O7WUX4_NEISU	O7WUX4 neisseria s
31	1007	69.6	280	1 LGTB_NEIGO	O50950 neisseria g

32	1003	69.4	280	2 O9REX4_NEIGO	O9REX4 neisseria g
33	1002.5	69.3	268	2 O7WUX1_NEIPO	O7WUX1 neisseria p
34	998	69.0	280	2 P96947_NEIME	P96947 neisseria m
35	996.5	68.9	268	2 O8K119_NEIME	O8K119 neisseria m
36	996	68.9	280	2 O5F4Y4_NEIG1	O5F4Y4 neisseria g
37	995.5	68.8	268	2 O8L2U8_NEIME	O8L2U8 neisseria m
38	969	67.0	280	2 O9RGM8_NEIGO	O9RGM8 neisseria g
39	962	66.5	280	2 O8L2V1_NEIME	O8L2V1 neisseria m
40	961	66.5	280	2 O8KR92_NEIME	O8KR92 neisseria m
41	951	65.8	280	2 O93EK4_NEIME	O93EK4 neisseria m
42	935	64.7	280	2 O9EVD3_NEISU	O9EVD3 neisseria s
43	933	64.5	280	1 LGTB_NEIME	O5117 neisseria m
44	932	64.5	280	2 O8L2V4_NEIME	O8L2V4 neisseria m
45	464	32.1	266	2 O8RKU0_HAEMIN	O8RKU0 haemophilus

ALIGNMENTS

RESULT 1

LGTB_NEIME STANDARD; PRT; 275 AA.

AC O5116;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Lacto-N-neotetraose biosynthesis glycosyl transferase 1gtb
 DE (EC 2.-.-.-).
 GN Name=1gtb; Ordered locus names=MNB1928;
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=96414473; PubMed=8817494;
 RA Jensen M.P., Hood D., Peak I.R.A., Virji M., Moxon E.R.;
 RT "Molecular analysis of a locus for the biosynthesis and phase-variable
 RT expression of the lacto-N-neotetraose terminal lipopolysaccharide
 RT structure in Neisseria meningitidis.";
 RL Mol. Microbiol. 18:729-740(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=2015755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
 RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
 RA Nelson K.E., Risen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
 RA Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R.T., Peterson J.D.,
 RA Hickley R.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
 RA Dougherty B.A., Mason T.M., Clecko A., Parksey D.S., Blair E.,
 RA Qian H., Vamathevan J.J., Gill J., Scarlato V., Masiatani V., Pizzo M.,
 RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappunli R.,
 RA Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58.";
 RL Science 287:1809-1815(2000).
 CC -1- FUNCTION: Adds the second galactose to the lacto-N-tetraose chain
 CC in lipopolysaccharide (LOS).
 CC -1- PATHWAY: Biosynthesis of the terminal lacto-N-neotetraose LPS
 CC structure.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 25 family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 DR EMBL; U25839; AAC44085.1; -; Genomic DNA.
 DR EMBL; AE002098; AAP42257.1; -; Genomic DNA.
 DR PIR; C81027; C81027.

DR PIR; S70814; S70814.
 DR TIGR; NMB1928; -.
 DR InterPro; IPR002654; Glyco_trans_25.
 DR PANTHER; PTHR15862; Glyco_trans_25; 1.
 DR Pfam; PF01755; Glyco_trans_25; 1.
 DR Complete proteome; Glycosyltransferase;
 KM Lipopolysaccharide biosynthesis; transferase.
 FT CONFLICT 97 97 A -> E (in Ref. 1).
 SQ SEQUENCE 275 AA; 31578 MW; E871305E2F6CF0F CRC64;

Query Match 99.6%; Score 1440; DB 1; Length 275;
 Best Local Similarity 99.6%; Pred. No. 2.2e-118;
 Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MONHVISLAAERRAHIAIDTGRGIGIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 DB 1 MONHVISLAAERRAHIAIDTGRGIGIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFDPDTAFIVR 120
 DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFDPDTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDKPAALPPEGGLHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDKPAALPPEGGLHP 180
 QY 181 VDLMMFSDFPREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 DB 181 VDLMMFSDFPREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 QY 241 PANTFKHRLIRALTISREKRRQRROQFIIVPQ 275
 DB 241 PANTFKHRLIRALTISREKRRQRROQFIIVPQ 275

RESULT 2

Q8KR92_NEIME PRELIMINARY; PRT; 275 AA.

AC Q8KR92;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Beta 1.4 galactosyltransferase.
 GN Name=lgTB;
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BB305-TT4;
 RK MEDLINE=22311028; PubMed=12423771; DOI=10.1016/S0928-8244(02)00379-6;
 RA Zhu P., Tsai C.M., Frasch C.E.;
 RT "Immunologic and genetic characterization of lipooligosaccharide
 RT variants in a Neisseria meningitidis serogroup C strain."
 RL FEMS Immunol. Med. Microbiol. 34:193-200(2002).
 DR EMBL; AY039684; XAK85139.1; -; Genomic DNA.
 DR GO; GO:0016757; F:transferase activity; transferring glycosyl. . .; IEA.
 DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR InterPro; IPR002654; Glyco_trans_25.
 DR Pfam; PF01755; Glyco_trans_25; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 275 AA; 31502 MW; 9C8CB6180A719A3B CRC64;

Query Match 98.7%; Score 1427; DB 2; Length 275;
 Best Local Similarity 98.5%; Pred. No. 3e-117;
 Matches 271; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MONHVISLAAERRAHIAIDTGRGIGIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 DB 1 MONHVISLAAERRAHIAIDTGRGIGIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFDPDTAFIVR 120

DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFDPDTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDKPAALPPEGGLHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDKPAALPPEGGLHP 180
 QY 181 VDLMMFSDFPREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 DB 181 VDLMMFSDFPREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 QY 241 PANTFKHRLIRALTISREKRRQRROQFIIVPQ 275
 DB 241 PANTFKHRLIRALTISREKRRQRROQFIIVPQ 275

RESULT 3

Q7WUX5_NEISU PRELIMINARY; PRT; 279 AA.

AC Q7WUX5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE lgTB.
 GN Name=lgTB;
 OS Neisseria subflava.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=28449;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Zhu P., Tsai C.M.;
 RT "Galactosyltransferase gene family in Neisseria: Implication for
 RT understanding the evolution of prokaryotic enzymes responsible for
 RT glycosylation."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY134877; AAN08515.1; -; Genomic DNA.
 SQ SEQUENCE 279 AA; 31861 MW; FDI8AA32E155A1 CRC64;

Query Match 96.7%; Score 1399; DB 2; Length 279;
 Best Local Similarity 96.4%; Pred. No. 9.1e-115;
 Matches 269; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 1 MONHVISLAAERRAHIAIDTGRGIGIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 DB 1 MONHVISLAAERRAHIAIDTGRGIGIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFDPDTAFIVR 120
 DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFDPDTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDKPAALPPEGGLHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDKPAALPPEGGLHP 180
 QY 181 VDLMMFSDFPREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 DB 181 VDLMMFSDFPREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 QY 241 PANTFKHRLIRALTISREKRRQRROQFIIVPQ 275
 DB 241 PANTFKHRLIRALTISREKRRQRROQFIIVPQ 279

RESULT 4

Q8L2V3_NEIME PRELIMINARY; PRT; 275 AA.

AC Q8L2V3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE lgTB.
 GN Name=lgTB;

```
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M992;
RX MEDLINE=22051050; PubMed=12055303;
RA Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three 1gt loci for biosynthesis of
  lipooligosaccharide (LOS) in Neisseria species.";
RL Microbiology 148:1833-1844(2002).
EMBL: AF470660; AM33865.1; -; Genomic DNA.
GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 275 AA; 31474 MW; 76C0BFA029DC9226 CRC64;

Query Match 96.3%; Score 1392; DB 2; Length 275;
Best Local Similarity 96.7%; Pred. No. 3.7e-114;
Matches 266; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MONHVISLASAERRAHADTGRGHPQFPDAMPSERLBOAMALVPGSAHPYLSG 60
DB 1 MONHVISLASAERRAHADTGRGHPQFPDAMPSERLBOAMALVPGSAHPYLSG 60
QY 61 VERACFMSHAIVMKQALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
DB 61 VERACFMSHAIVMKQALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
QY 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
DB 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
QY 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
DB 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
QY 181 VDLMMSPDFDREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNKKOQRDS 240
DB 181 VDLMMSPDFDREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNKKOQRDS 240
QY 241 PANTFGRRLIRALTITKISRERKRRQRROQFIVPQ 275
DB 241 PANTFGRRLIRALTITKISRERKRRQRROQFIVPQ 275

RESULT 5
Q8L2V5 NEIME PRELIMINARY; PRT; 275 AA.
ID Q8L2V5;
AC Q8L2V5;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE LgtB.
GN Name=LgtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M981;
RX MEDLINE=22051050; PubMed=12055303;
RA Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three 1gt loci for biosynthesis of
  lipooligosaccharide (LOS) in Neisseria species.";
RL Microbiology 148:1833-1844(2002).
EMBL: AA470659; AM33862.1; -; Genomic DNA.
GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 275 AA; 31543 MW; 6EB00EB09A128DE CRC64;

Query Match 96.2%; Score 1391; DB 2; Length 275;
Best Local Similarity 96.7%; Pred. No. 4.5e-114;
Matches 266; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 1 MONHVISLASAERRAHADTGRGHPQFPDAMPSERLBOAMALVPGSAHPYLSG 60
DB 1 MONHVISLASAERRAHADTGRGHPQFPDAMPSERLBOAMALVPGSAHPYLSG 60
QY 61 VERACFMSHAIVMKQALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
DB 61 VERACFMSHAIVMKQALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
QY 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
DB 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
QY 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
DB 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
QY 181 VDLMMSPDFDREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNKKOQRDS 240
DB 181 VDLMMSPDFDREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNKKOQRDS 240
QY 241 PANTFGRRLIRALTITKISRERKRRQRROQFIVPQ 275
DB 241 PANTFGRRLIRALTITKISRERKRRQRROQFIVPQ 275

RESULT 6
Q8L2V7 NEIME PRELIMINARY; PRT; 275 AA.
ID Q8L2V7;
AC Q8L2V7;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE LgtB.
GN Name=LgtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=891;
RX MEDLINE=22051050; PubMed=12055303;
RA Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three 1gt loci for biosynthesis of
  lipooligosaccharide (LOS) in Neisseria species.";
RL Microbiology 148:1833-1844(2002).
EMBL: AF470658; AM33859.1; -; Genomic DNA.
GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 275 AA; 31530 MW; 5C9BBD41BAAB04F CRC64;

Query Match 95.9%; Score 1386; DB 2; Length 275;
Best Local Similarity 96.4%; Pred. No. 1.2e-113;
Matches 265; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MONHVISLASAERRAHADTGRGHPQFPDAMPSERLBOAMALVPGSAHPYLSG 60
DB 1 MONHVISLASAERRAHADTGRGHPQFPDAMPSERLBOAMALVPGSAHPYLSG 60
QY 61 VERACFMSHAIVMKQALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
DB 61 VERACFMSHAIVMKQALDEGLPYITVFEEDVLLGEGEKEFLADAWLQERFPDPTAFIVR 120
QY 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
DB 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
QY 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
DB 121 LETMFHNVLTSSGVAADYCGRAFPILSESHMGTAGYIISRKMRFFLDRFALPPEGGLHP 180
QY 181 VDLMMSPDFDREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNKKOQRDS 240
DB 181 VDLMMSPDFDREGMPVCOINPALCAQELHYAKFHDONSALGSLIEHRLNKKOQRDS 240
QY 241 PANTFGRRLIRALTITKISRERKRRQRROQFIVPQ 275
DB 241 PANTFGRRLIRALTITKISRERKRRQRROQFIVPQ 275
```



```

RESULT 7
Q9EVD4 NEISU
ID Q9EVD4 NEISU PRELIMINARY; PRT; 275 AA.
AC Q9EVD4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Galactosyl transferase.
GN Name=1gtB;
OS Neisseria subflava.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=28449;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21142520; PubMed=11208792; DOI=10.1128/JB.183.3.934-941.2001;
RT Arking D., Tong Y., Stein D.C.;
RT "Analysis of lipooligosaccharide biosynthesis in the Neisseriaceae.";
RL J. Bacteriol. 183:934-941(2001).
DR EMBL; AF240672; AAC09765.1; -; Genomic DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR002654; Glyco_trans_25.
DR Pfam; PF01755; Glyco_transf_25; 1.
KW Transferase.
SQ SEQUENCE 275 AA; 31441 MW; D719F3B15F64D14C CRC64;

Query Match 94.1%; Score 1361; DB 2; Length 275;
Best Local Similarity 94.9%; Pred. No. 2e-111;
Matches 261; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MONHVISLASAARRAHADTFGRHGIPQFPDMLMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLASAARRAHADTFGRHGIPQFPDMLMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VERACFMSHAVALMKQALDEGLPYITVFEDDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
DB 61 VERACFMSHAVALMKQALDEGLPYITVFEDDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMFMAVLTSPSGVADYCGRAFPLLSESHMGTAGYIISRKAMRFLDRFALPPEGGLHP 180
DB 121 LETMFMAVLTSPSGVADYCGRAFPLLSESHMGTAGYIISRKAMRFLDRFALPPEGGLHP 180
QY 181 VDLMMFSDFPDRBEGMPVCQALNPALCAQELHYAKFHONSAUGSLIEHDLINRKQQRDS 240
DB 181 VDLMMFSDFPDRBEGMPVCQALNPALCAQELHYAKFHONSAUGSLIEHDLINRKQQRDS 240
QY 241 PANTFKHRLIRALTISRERERKRRORREOFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERERKRRORREOFIVPFQ 275

RESULT 8
Q8K1A8 NEIME
ID Q8K1A8 NEIME PRELIMINARY; PRT; 279 AA.
AC Q8K1A8;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE LgtB.
GN Name=1gtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=M120; and 7880;
RX MEDLINE=22051050; PubMed=12055303;
RX Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three lgt loci for biosynthesis of
lipooligosaccharide (LOS) in Neisseria species.";

```

```

RL Microbiology 148:1833-1844(2002).
DR EMBL; AF470662; AAM3870.1; -; Genomic DNA.
DR EMBL; AF470663; AAM3872.1; -; Genomic DNA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR002654; Glyco_trans_25.
DR Pfam; PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 279 AA; 31660 MW; 099D2E1E0DF456F6 CRC64;

Query Match 94.1%; Score 1360; DB 2; Length 279;
Best Local Similarity 94.6%; Pred. No. 2.5e-111;
Matches 264; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 1 MONHVISLASAARRAHADTFGRHGIPQFPDMLMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLASAARRAHADTFGRHGIPQFPDMLMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VERACFMSHAVALMKQALDEGLPYITVFEDDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
DB 61 VERACFMSHAVALMKQALDEGLPYITVFEDDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMFMAVLTSPSGVADYCGRAFPLLSESHMGTAGYIISRKAMRFLDRFALPPEGGLHP 180
DB 121 LETMFMAVLTSPSGVADYCGRAFPLLSESHMGTAGYIISRKAMRFLDRFALPPEGGLHP 180
QY 181 VDLMMFSDFPDRBEGMPVCQALNPALCAQELHYAKFHONSAUGSLIEHDLINRKQQRDS 240
DB 181 VDLMMFSDFPDRBEGMPVCQALNPALCAQELHYAKFHONSAUGSLIEHDLINRKQQRDS 240
QY 241 PANTFKHRLIRALTISRERERKRRORREOFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERERKRRORREOFIVPFQ 275

RESULT 9
Q7WUX2 NEIPO
ID Q7WUX2 NEIPO PRELIMINARY; PRT; 279 AA.
AC Q7WUX2;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE LgtB.
GN Name=1gtB;
OS Neisseria polysacchara.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=489;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Zhu P., Tsai C.M.;
RT "Galactosyltransferase gene family in Neisseria: Implication for
understanding the evolution of prokaryotic enzymes responsible for
RT glycosylation.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF134878; AAN08518.1; -; Genomic DNA.
DR EMBL; AF134878; AAN08518.1; -; Genomic DNA.
SQ SEQUENCE 279 AA; 31907 MW; 60BDAPBD749D46D5 CRC64;

Query Match 93.6%; Score 1353; DB 2; Length 279;
Best Local Similarity 93.5%; Pred. No. 1e-110;
Matches 261; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 1 MONHVISLASAARRAHADTFGRHGIPQFPDMLMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLASAARRAHADTFGRHGIPQFPDMLMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VERACFMSHAVALMKQALDEGLPYITVFEDDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
DB 61 VERACFMSHAVALMKQALDEGLPYITVFEDDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMFMAVLTSPSGVADYCGRAFPLLSESHMGTAGYIISRKAMRFLDRFALPPEGGLHP 180
DB 121 LETMFMAVLTSPSGVADYCGRAFPLLSESHMGTAGYIISRKAMRFLDRFALPPEGGLHP 180
QY 181 VDLMMFSDFPDRBEGMPVCQALNPALCAQELHYAKFHONSAUGSLIEHDLINRKQQRDS 240

```



```

Db      181 VDMWNGNDPDRGMPVCOINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
      241 PANTFKHRLIRALTISRERKRRORREO----FIVPQ 275
      241 PANTFKHRLIRALTISRERKRRORREOIKGKTIVPQ 279

RESULT 10
OBL2V9 NEIME
ID      OBL2V9_NEIME PRELIMINARY; PRT; 279 AA.
AC      OBL2V9;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DB      LGTB.
GN      Name=lgTB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX      Neisseriaceae; Neisseria.
NCBI_TaxID=487;
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=35E;
RX      MEDLINE=22051050; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
      lipooligosaccharide (LOS) in Neisseria species.";
RL      Microbiology 148:1833-1844(2002).
EMBL; AF470655; AM33850.1; -; Genomic DNA.
GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco trans 25; I.
DR      Pfam; PF01755; Glyco trans 25; I.
SQ      SEQUENCE 279 AA; 31962 MW; 4321C9A18C44D075 CRC64;

Query Match      92.2%; Score 1333; DB 2; Length 279;
Best Local Similarity 92.1%; Pred. No. 5.9e-109;
Matches 257; Conservative 6; Mismatches 12; Indels 4; Gaps 1;

QY      1 MONHVISLASAERRAHIAIDTFRGRIPOFPDAMPSERLQAMALVPGLSAHPTLSG 60
      1 MONHVISLASAERRAHIAIDTFRGRIPOFPDAMPSERLQAMALVPGLSAHPTLSG 60
DB      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
DB      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
QY      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
DB      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
QY      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
DB      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
QY      241 PANTFKHRLIRALTISRERKRRORREO----FIVPQ 275
      241 PANTFKHRLIRALTISRERKRRORREOIKGKTIVPQ 279
DB      241 PANTFKHRLIRALTISRERKRRORREOIKGKTIVPQ 279

RESULT 11
OBL2V9 NEIME PRELIMINARY; PRT; 275 AA.
AC      OBL2V9;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT      01-FEB-2005 (TEMBLrel. 29, Last annotation update)
DB      LGTB.
GN      Name=lgTB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX      Neisseriaceae; Neisseria.
NCBI_TaxID=487;

```

```

RN      [1]
RC      NUCLEOTIDE SEQUENCE.
RC      STRAIN=6275, and M986;
RX      MEDLINE=22051050; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
      lipooligosaccharide (LOS) in Neisseria species.";
RL      Microbiology 148:1833-1844(2002).
EMBL; AF470655; AM33850.1; -; Genomic DNA.
GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco trans 25; I.
DR      Pfam; PF01755; Glyco trans 25; I.
SQ      SEQUENCE 275 AA; 31558 MW; A73037B079FF5DB3 CRC64;

Query Match      91.8%; Score 1328; DB 2; Length 275;
Best Local Similarity 92.0%; Pred. No. 1.6e-108;
Matches 253; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY      1 MONHVISLASAERRAHIAIDTFRGRIPOFPDAMPSERLQAMALVPGLSAHPTLSG 60
      1 MONHVISLASAERRAHIAIDTFRGRIPOFPDAMPSERLQAMALVPGLSAHPTLSG 60
DB      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
DB      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
      61 VERACFMSHAVLMKQALDSEGLPYITVFEDVLLGEGEKFLEADAWLQSRFPDPTAFYR 120
QY      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
DB      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAAGTISRKMRPFLDRFPAALPPGGLHP 180
QY      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
DB      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
      181 VDLMFSDPFDEMGVPCQINPALCAQELHYAKFHONSALGSLIHDRLNKKQRDS 240
QY      241 PANTFKHRLIRALTISRERKRRORREOIFIVPQ 275
      241 PANTFKHRLIRALTISRERKRRORREOIFIVPQ 275
DB      241 PANTFKHRLIRALTISRERKRRORREOIFIVPQ 275

RESULT 12
OBL2V2 NEIME PRELIMINARY; PRT; 275 AA.
ID      OBL2V2_NEIME PRELIMINARY; PRT; 275 AA.
AC      OBL2V2;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DB      LGTB.
GN      Name=lgTB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX      Neisseriaceae; Neisseria.
NCBI_TaxID=487;
RN      [1]
RC      NUCLEOTIDE SEQUENCE.
RC      STRAIN=6155;
RX      MEDLINE=22051050; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
      lipooligosaccharide (LOS) in Neisseria species.";
RL      Microbiology 148:1833-1844(2002).
EMBL; AF470661; AM33867.1; -; Genomic DNA.
GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco trans 25; I.
DR      Pfam; PF01755; Glyco trans 25; I.
SQ      SEQUENCE 275 AA; 31623 MW; B435413181C35674 CRC64;

Query Match      91.8%; Score 1327; DB 2; Length 275;
Best Local Similarity 92.4%; Pred. No. 2e-108;
Matches 254; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY      1 MONHVISLASAERRAHIAIDTFRGRIPOFPDAMPSERLQAMALVPGLSAHPTLSG 60
      1 MONHVISLASAERRAHIAIDTFRGRIPOFPDAMPSERLQAMALVPGLSAHPTLSG 60

```

```

Db      1  M0NHV1SLASAARRAHIAATGARGIPFOFPDAMPSERLQAMAEVPGISAHPIYLSG 60
Qy      61  VAKACFMSHAVALMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLQERPPDTAFYR 120
Db      61  VAKACFMSHAVALMKQALDEGLPYIAVFEEDVLLGEGEKFLEADAWLKERPPDPSAFYR 120
Qy      121  LETMFMHVLTPSPGVADYCGRAFPILLESHEMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Db      121  LETMFMHVLTPSPGVADYCGRAFPILLESHEMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Qy      181  VDLMMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNRKQQRDS 240
Db      181  VDLMMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNRKQQRDS 240
Qy      241  PANTFKHRLIRALTISREERKRORREOFIYVFPQ 275
Db      241  PANTFKHRLIRALTISREERKRORREOFIYVFPQ 275

```

```

RESULT 13
08L2U6_NEIME
ID      08L2U6_NEIME PRELIMINARY; PRT; 279 AA.
AC      08L2U6;
DT      01-OCT-2002 (T#BMBLrel. 22, Created)
DT      01-OCT-2002 (T#EMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (T#EMBLrel. 24, Last annotation update)
DE      LGTB.
GN      Name=lgTB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC      Neisseriaceae; Neisseria.
OX      NCBI_TaxID=487;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=22051050; PubMed=12055303;
RA      Zhu P.; Klutch M.J.; Bash M.C.; Tsang R.S.W.; Ng L.K.; Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
RL      lipooligosaccharide (LOS) in Neisseria species.";
RL      Microbiology 148:1833-1844(2002).
DR      EMBL; AF470665; AM33876.1; -; Genomic DNA.
DR      GO; GO:0009103; P.lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco_trans_25.
DR      Pfam; PF01755; Glyco_trans_25; 1.
SQ      SEQUENCE 279 AA; 31927 MW; C7D3A3BE74ABA3D9 CRC64;

```

```

Query Match      91.6%; Score 1325; DB 2; Length 279;
Best Local Similarity 92.5%; Pred. No. 3e-108;
Matches 258; Conservative 3; Mismatches 14; Indels 4; Gaps 1;

Qy      1  M0NHV1SLASAARRAHIAATGARGIPFOFPDAMPSERLQAMAEVPGISAHPIYLSG 60
Db      1  M0NHV1SLASAARRAHIAATGARGIPFOFPDAMPSERLQAMAEVPGISAHPIYLSG 60
Qy      61  VAKACFMSHAVALMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLQERPPDTAFYR 120
Db      61  VAKACFMSHAVALMKQALDEGLPYIAVFEEDVLLGEGEKFLEADAWLQERPPDPSAFYR 120
Qy      121  LETMFMHVLTPSPGVADYCGRAFPILLESHEMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Db      121  LETMFMHVLTPSPGVADYCGRAFPILLESHEMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Qy      181  VDLMMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNRKQQRDS 240
Db      181  VDLMMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNRKQQRDS 240
Qy      241  PANTFKHRLIRALTISREERKRORREOFIYVFPQ 275
Db      241  PANTFKHRLIRALTISREERKRORREOFIYVFPQ 275

```

RESULT 14
LGTB_NEIME

```

ID      LGTB_NEIME STANDARD; PRT; 279 AA.
AC      P57033;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      13-SEP-2005 (Rel. 48, Last annotation update)
DE      Lacto-N-neotetraose biosynthesis glycosyl transferase lgtB
DE      (EC 2.4.1.21).
GN      Name=lgTB; OrderedLocustNames=MMA0525;
OS      Neisseria meningitidis (serogroup A).
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC      Neisseriaceae; Neisseria.
OX      NCBI_TaxID=65699;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX      STRAIN=22491 / Serogroup A / Serotype 4a;
RX      MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA      Parthill J., Achman M., James K.D., Bentley S.D., Churcher C.M.,
RA      Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA      Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA      Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA      Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA      Whitehead S., Spratt B.G., Barrall B.G.;
RT      "Complete DNA sequence of a serogroup A strain of Neisseria
RL      meningitidis 22491.";
RL      Nature 404:502-506(2000).
CC      -1- FUNCTION: Adds the second galactose to the lacto-N-tetraose chain
CC      in lipooligosaccharide (LOS).
CC      -1- PATHWAY: Biosynthesis of the terminal lacto-N-neotetraose LPS
CC      structure.
CC      -1- SIMILARITY: Belongs to the glycosyltransferase 25 family.
CC      CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      CC
CC      DR      EMBL; AL162753; CAB83817.1; -; Genomic DNA.
CC      DR      PIR; A81971; A81971.
CC      DR      InterPro; IPR002654; Glyco_trans_25.
CC      DR      PANTHER; PTHR15862; Glyco_trans_25; 1.
CC      DR      Pfam; PF01755; Glyco_trans_25; 1.
CC      KW      Complete proteome; Glycosyltransferase;
CC      KW      Lipopolysaccharide biosynthesis; Transferase.
SQ      SEQUENCE 279 AA; 31904 MW; 8703B56513A0D347 CRC64;

```

```

Query Match      91.6%; Score 1324; DB 1; Length 279;
Best Local Similarity 92.5%; Pred. No. 3.7e-108;
Matches 258; Conservative 2; Mismatches 15; Indels 4; Gaps 1;

Qy      1  M0NHV1SLASAARRAHIAATGARGIPFOFPDAMPSERLQAMAEVPGISAHPIYLSG 60
Db      1  M0NHV1SLASAARRAHIAATGARGIPFOFPDAMPSERLQAMAEVPGISAHPIYLSG 60
Qy      61  VAKACFMSHAVALMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLQERPPDTAFYR 120
Db      61  VAKACFMSHAVALMKQALDEGLPYIAVFEEDVLLGEGEKFLEADAWLQERPPDPSAFYR 120
Qy      121  LETMFMHVLTPSPGVADYCGRAFPILLESHEMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Db      121  LETMFMHVLTPSPGVADYCGRAFPILLESHEMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Qy      181  VDLMMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNRKQQRDS 240
Db      181  VDLMMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNRKQQRDS 240
Qy      241  PANTFKHRLIRALTISREERKRORREOFIYVFPQ 275
Db      241  PANTFKHRLIRALTISREERKRORREOFIYVFPQ 275

```

RESULT 15
Q93EK5_NEIME

ID O93BK5_NEIME_PRELIMINARY; PRT; 275 AA.
AC O93BK5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LgtB.
GN Name=LgtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M978;
RX MEDLINE=21467954; Pubmed=11583844;
RA Zhu P., Klutch M.J., Tsai C.-M.;
RT "Genetic analysis of conservation and variation of lipooligosaccharide
expression in two L8-immunotype strains of Neisseria meningitidis.";
RL FEMS Microbiol. Lett. 203:173-177(2001).
DR EMBL: AF355193; AAL12841.1; -; Genomic DNA.
DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; I.
SQ SEQUENCE 275AA; 31577 MW; A73037E0625EB7B3 CRC64;

Query Match 91.3%; Score 1320; DB 2; Length 275;
Best Local Similarity 91.6%; Pred. No. 8,1e-108;
Matches 252; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 MGNHVISLSAARRAHIAADTPGRHGIPOFPDAMPSERLBOAMALVPGLSAHPYLSG 60
DB 1 MGNHVISLSAARRAHIAATFGARGIPPOFPDAMPSERLBOAMALVPGLSAHPYLSG 60
QY 61 VERACMSNAVLKQALDEGLPYITVFEDVLLGEGEKEFLAEDAMLOERFPDPTAFIVR 120
DB 61 VERACMSNAVLKQALDEGLPYIAVFEDVLLGEGEKEFLAEDAMLOERFPDPTAFIVR 120
QY 121 LETMFHVLTPSGVADYCGRAFPPLSEHMGTAGYTIISRKAMPPLDRPALPPGLHP 180
DB 121 LETMFHVLTPSGVADYCGRAFPPLSEHMGTAGYTIISRKAMPPLDRPALPPGLHP 180
QY 181 VDLMFSDPFDRBGMPCQALPALCAQELHYAKFDONSALGSLIEHRLNKKQQRDS 240
DB 181 VDLMFSDPFDRBGMPCQALPALCAQELHYAKFDONSALGSLIEHRLNKKQQRDS 240
QY 241 PANTFKRLIRALTISRERKRRORREOPTVPPQ 275
DB 241 PANTFKRLIRALTISRERKRRORREOPTVPPQ 275

Search completed: April 7, 2006, 15:29:48
Job time : 252 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2006, 15:30:06 ; Search time 47 Seconds
(without alignments)
483.741 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHYISLASAERRAHAD.....ISREKRKRQRREQIVPQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5 COMB.pep:*
2: /cgn2_6/prodata/1/1aa/6 COMB.pep:*
3: /cgn2_6/prodata/1/1aa/7 COMB.pep:*
4: /cgn2_6/prodata/1/1aa/8 COMB.pep:*
5: /cgn2_6/prodata/1/1aa/9 COMB.pep:*
6: /cgn2_6/prodata/1/1aa/10 COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1288	89.1	279	1	US-08-312-387B-8
2	1288	89.1	279	1	US-08-683-426-8
3	1288	89.1	279	1	US-08-683-458-8
4	1288	89.1	279	1	US-08-878-360-8
5	1288	89.1	279	2	US-09-333-412-8
6	1288	89.1	279	2	US-10-007-267A-8
7	1007	69.6	280	1	US-08-312-387B-6
8	1007	69.6	280	1	US-08-683-426-6
9	1007	69.6	280	1	US-08-683-458-6
10	1007	69.6	280	1	US-08-878-360-6
11	1007	69.6	280	2	US-08-478-140B-6
12	1007	69.6	280	2	US-09-333-412-6
13	1007	69.6	280	2	US-09-338-943-6
14	1007	69.6	280	2	US-10-007-267A-6
15	364	25.2	255	2	US-09-540-236-2407
16	118.5	8.2	579	2	US-09-369-247-61
17	118.5	8.2	579	2	US-10-062-548-61
18	98.5	6.8	273	2	US-09-601-519A-1
19	98.5	6.8	273	2	US-10-218-381-1
20	93	6.4	114	2	US-09-344-587-10
21	93	6.4	150	1	US-08-400-115-2
22	93	6.4	271	1	US-08-400-115-4
23	92	6.4	244	2	US-09-244-369B-1
24	92	6.4	244	2	US-09-940-391-1
25	92	6.4	255	2	US-09-328-352-5253
26	91	6.3	116	2	US-09-244-592-1
27	91	6.3	133	2	US-08-961-309-48

28	90	6.2	5069	2	US-10-042-665A-5	Sequence 5, Appl1
29	89	6.2	133	1	US-08-822-028-6	Sequence 6, Appl1
30	89	6.2	133	1	US-08-822-028-30	Sequence 30, Appl1
31	89	6.2	133	2	US-08-479-285-6	Sequence 6, Appl1
32	89	6.2	133	2	US-08-479-285-30	Sequence 30, Appl1
33	89	6.2	133	2	US-09-503-653A-6	Sequence 6, Appl1
34	89	6.2	133	2	US-09-503-653A-30	Sequence 30, Appl1
35	88.5	6.1	134	2	US-08-961-309-49	Sequence 49, Appl1
36	87.5	6.1	238	1	US-08-392-338A-12	Sequence 12, Appl1
37	87.5	6.1	238	1	US-08-926-789-12	Sequence 12, Appl1
38	87.5	6.1	238	2	US-09-166-750-21	Sequence 21, Appl1
39	87.5	6.1	238	2	US-09-166-093-21	Sequence 21, Appl1
40	87.5	6.1	238	2	US-09-172-019-21	Sequence 21, Appl1
41	87.5	6.1	238	2	US-09-166-094-21	Sequence 21, Appl1
42	87.5	6.1	238	2	US-09-443-213-21	Sequence 21, Appl1
43	87.5	6.1	239	4	PCT-US93-11138-12	Sequence 12, Appl1
44	87.5	6.1	134	1	US-08-822-028-10	Sequence 10, Appl1
45	86.5	6.0				

ALIGNMENTS

```
RESULT 1
US-08-312-387B-8
; Sequence 8, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-312-387B-8

Query Match      89.1%; Score 1288; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 9.9e+143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
```

QY 1 MGNHYISLASAERRAHADTFGRHGIPFPFDALMPSERLRQMAELVPGISAPHYISG 60
DB 1 MGNHYISLASAERRAHADTFGRHGIPFPFDALMPSERLRQMAELVPGISAPHYISG 60
QY 61 VEKACFMHSAVLKQALDEGLPYITVFEDDVILGEGEKKFLAEDAMLDGERPDPAFTYR 120

DB 61 VAKACFMSHVAVLMQALDEGVYIAVFEEDVLLGSGAEQFLAEDTWLOERFPDPSAFVVR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSEHMTAGYIISRKAMRFFLDLRFALPPEGILHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSEHMTAGYIISRKAMRFFLDLRFALPPEGILHP 180
QY 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDQNSALGSLIEHDLNRKQORRDS 240
DB 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDQNSALGSLIEHDLNRKQORRDS 240
QY 241 PANTFKHRLIRALTITKISREKRRQRREQ----FIVPFO 275
DB 241 PANTFKHRLIRALTITKISREKRRQRREQ----FIVPFO 275

RESULT 2
US-08-683-426-8
; Sequence 8, Application US/08683426
; Patent No. 5705367
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,426
; FILING DATE:
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-426-8

Query Match 89.1%; Score 1288; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 9,9e-143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
QY 1 MONHVISLASAERRAHIAITFGSRGIPQFDPALMPSSRLERAAELVPGLSAHPLYSG 60
DB 1 MONHVISLASAERRAHIAITFGSRGIPQFDPALMPSSRLERAAELVPGLSAHPLYSG 60
QY 61 VAKACFMSHVAVLMQALDEGLPYITVFEEDVLLGSGAEKFLAEDAWLOERFPDPTAFIVR 120
DB 61 VAKACFMSHVAVLMQALDEGLPYITVFEEDVLLGSGAEKFLAEDAWLOERFPDPTAFIVR 120

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSEHMTAGYIISRKAMRFFLDLRFALPPEGILHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSEHMTAGYIISRKAMRFFLDLRFALPPEGILHP 180
QY 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDQNSALGSLIEHDLNRKQORRDS 240
DB 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDQNSALGSLIEHDLNRKQORRDS 240
QY 241 PANTFKHRLIRALTITKISREKRRQRREQ----FIVPFO 275
DB 241 PANTFKHRLIRALTITKISREKRRQRREQ----FIVPFO 275

RESULT 3
US-08-683-458-8
; Sequence 8, Application US/08683458
; Patent No. 5798233
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,458
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-458-8

Query Match 89.1%; Score 1288; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 9,9e-143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
QY 1 MONHVISLASAERRAHIAITFGSRGIPQFDPALMPSSRLERAAELVPGLSAHPLYSG 60
DB 1 MONHVISLASAERRAHIAITFGSRGIPQFDPALMPSSRLERAAELVPGLSAHPLYSG 60
QY 61 VAKACFMSHVAVLMQALDEGLPYITVFEEDVLLGSGAEKFLAEDAWLOERFPDPTAFIVR 120
DB 61 VAKACFMSHVAVLMQALDEGLPYITVFEEDVLLGSGAEKFLAEDAWLOERFPDPTAFIVR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSEHMTAGYIISRKAMRFFLDLRFALPPEGILHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSEHMTAGYIISRKAMRFFLDLRFALPPEGILHP 180

QY	181	VDIAMEFSDPEDEGMPVCCOLNPLCAQOEIHYAFHONSALGSLIEHDLRLNKKOORDS	240
Db	181	VDIAMEFGNDDEGMPVCCOLNPLCAQOEIHYAFHONSALGSLIEHDLRLNKKOORDS	240
QY	241	PANTFKHRLIRALTYKISREKERKORREO----	275
Db	241	PANTFKHRLIRALTYKISREKERKORREOLIGKIIVPFO	279

RESULT 4

US-08-878-360-8
; Sequence 8, Application US/08878360
; Patent No. 5945322

GENERAL INFORMATION:

1 APPLICANT: Gotschlich, Emil C.
2
3 TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
4
5 TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
6
7 NUMBER OF SEQUENCES: 12
8
9 CORRESPONDENCE ADDRESS:
10
11 ADDRESSER: Klaber & Jackson
12
13 STREET: 411, Hackensack Avenue
14
15 CITY: Hackensack
16
17 STATE: New Jersey
18
19 COUNTRY: USA

COMPUTER READABLE FORM:

1 MEDIUM TYPE: Floppy disk
 2 COMPUTER: IBM PC compatible
 3 OPERATING SYSTEM: PC-DOS/MS-DOS
 4 SOFTWARE: Patenclin Release #1.0, Version #1.25
 5
 6 CURRENT APPLICATION DATA:
 7 APPLICATION NUMBER: US/08/878,360
 8 FILING DATE: 18-JUN-1997

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/683,426
; FILING DATE:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 199
;

```

ATTORNEY/AGENT INFORMATION

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:

TELEFAX: 201 343-1688

TELEX: 133521

; INFORMATION FOR SEQ ID NO: 8:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid

```

MOLECULE TYPE: protein

US-08-878-360-8

Query Match

Best Local Similarity 90.0%; Pred. No. 9.9e-143;
Matches 251; Conservative 7; Mismatches 17; Indels

[illegible]

QY	181	VDLMMFSQFPREGPVCQINLPALCAQELHAYKFPDQNSALSLIENDRLNNKQQRDS	240
Db	181	VDLMMFGNPDREGPVCQINLPALCAQELHAYKFPDQNSALSLIENDRLNNKQQRDS	240
QY	241	PANTFKHLIRALTYKISREKRRQRREC----	PTVFFQ 275
Db	241	PANTFKHLIRALTYKIGRERGRRRRQQLGIITVFFQ	279

RESULT 5

US-09-333-412-8
; Sequence 8, Application US/09333412
; Patent No. 6342382

GENERAL INFORMATION:

? APPLICANT: GÖRGENSCHILD, EMIL C.
 ? TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
 ? OLIGOSACCHARIDES, AND GENES ENCODING THEM
 ?
 ? NUMBER OF SEQUENCES: 12
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSER: Klauber & Jackson
 ? STREET: 411 Hackensack Avenue
 ? CITY: Hackensack
 ? STATE: New Jersey
 ? COUNTRY: USA
 ? ZIP: 07601
 ?

MEDIUM TYPE: FLOPPY

```

?
?      COMPUTER: IBM PC compatible
?
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?
?      SOFTWARE: Patentin Release #1.0, Version #1.25
?
?      CURRENT APPLICATION DATA:
?
?      APPLICATION NUMBER: US/09/333,412
?
?      FILING DATE: 15-Jun-1999
?
?      CLASSIFICATION: <Unknown>
?

```

APPLICATION NUMBER

1 FILING DATE: July 7, 1994
2
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Jackson Esq., David A.
5 REGISTRATION NUMBER: 26,742
6

TELECOMMUNICATION INFO

TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8

LENGTH: 279 amino

TYPE: amino acid

TOPOLOGY: linear

```

;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-333-412-8

```

Query Match

Best Local SI

1 MONHVISLASAERRAHITDTFGRHGI PFOFPDAMPSERL EQAMARLVPQ

Db	1	MÖNNVISAASAERRAHIAATFGSGKGIPOFFDADAMPSEBRLERMAAEIVPGLSAPHYLSG	60
Qy	61	VEKACFMSHATLMTQALDEGPIYITVFPPDDVLLGSEBKEFTLADAWLQRPDPDTAFYR	120
Db	61	VEKACFMSHATLMTQALDEGPIYIAVFPPDDVLLGSGAQPFLAEDPTWLQRPDPDSAFYR	120
Qy	121	LETMPMYLTSPSGVADYCGRAFLPLBSEHMGTAQYIISRKMRFFLDRFALPBGGLHP	180
Db	121	LETMPMYLTSPSGVADYCGRAFLPLBSEHGTAGYIISRKMRFFLDRFAVLPBRLHP	180
Qy	181	VDLMMESDFPREGMPVQVQLPALCAQELHYAKFHDQNSALGSLTEHRLRLNRKQQRSS	240
Db	181	VDLMMESGPDREGMPVQVQLPALCAQELHYAKFHDQNSALGSLTEHRLRLNRKQQRSS	240

QY 241 PANTFKHRLIRALTKISREKRRORRQ----FIVPQ 275
DB 241 PANTFKHRLIRALTKISREKRRORRQELIGKIIVPQ 279

RESULT 6
US-10-007-267A-8
Sequence 8, Application US/10007267A
Patent No. 6780624

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
FILE REFERENCE: 040853-01-5029-02
CURRENT APPLICATION NUMBER: US/10/007,267A
CURRENT FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US 09/333,412
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/878,360
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: US 08/683,426
PRIOR FILING DATE: 1996-07-18
PRIOR APPLICATION NUMBER: US 08/312,387
PRIOR FILING DATE: 1994-09-24
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin version 3.2
SEQ ID NO 8
LENGTH: 279
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-007-267A-8

Query Match 89.1%; Score 1288; DB 2; Length 279;
Best Local Similarity 90.0%; Pred. No. 9,9e-143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVISLSAASARRAHADTFGRHGIPQFPDMLPSERLQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLSAASARRAHADTFGRHGIPQFPDMLPSERLQMAELVPGISAHPIYLSG 60
QY 61 VKACPMASHAVIMKQALDDEGLPIYTVFEDDVLGSGEKKFLAEDAMLOERPDPTAFIVR 120
DB 61 VKACPMASHAVIMKQALDDEGLPIYTVFEDDVLGSGEKKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMEHVLITSPSGVADYCGRAFPILSEHMGWAGYIISRKMRFFLDRAALPPEGILHP 180
DB 121 LETMEHVLITSPSGVADYCGRAFPILSEHMGWAGYIISRKMRFFLDRAALPPEGILHP 180
QY 181 VDLMMFSDPDEBGMVQCLNPLCAQELHYAKFHONSALGSLIETHDLNLRKQORRS 240
DB 181 VDLMMFSDPDEBGMVQCLNPLCAQELHYAKFHONSALGSLIETHDLNLRKQORRS 240
QY 241 PANTFKHRLIRALTKISREKRRORRQ----FIVPQ 275
DB 241 PANTFKHRLIRALTKISREKRRORRQELIGKIIVPQ 279

RESULT 7
US-08-312-387B-6
Sequence 6, Application US/08312387B
Patent No. 5545553
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA

ZIP: 07601
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/312,387B
FILING DATE: July 7, 1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 280 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-312-387B-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVISLSAASARRAHADTFGRHGIPQFPDMLPSERLQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLSAASARRAHADTFGRHGIPQFPDMLPSERLQMAELVPGISAHPIYLSG 60
QY 61 VKACPMASHAVIMKQALDDEGLPIYTVFEDDVLGSGEKKFLAEDAMLOERPDPTAFIVR 120
DB 61 VKACPMASHAVIMKQALDDEGLPIYTVFEDDVLGSGEKKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMEHVLITSPSGVADYCGRAFPILSEHMGWAGYIISRKMRFFLDRAALPPEGILHP 180
DB 121 LETMEHVLITSPSGVADYCGRAFPILSEHMGWAGYIISRKMRFFLDRAALPPEGILHP 180
QY 181 VDLMMFSDPDEBGMVQCLNPLCAQELHYAKFHONSALGSLIETHDLNLRKQORRS 240
DB 181 VDLMMFSDPDEBGMVQCLNPLCAQELHYAKFHONSALGSLIETHDLNLRKQORRS 240
QY 241 PANTFKHRLIRALTKISREKRRORRQ 269
DB 241 PANTFKHRLIRALTKISREKRRORRQ 265

RESULT 8
US-08-683-426-6
Sequence 6, Application US/08683426
Patent No. 5705367
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/683,426
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-426-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACMSHAIVLMKQALDGLPYITVFEDVLLGEGEKEFLADAWLQERFPDPTAFYR 120
DB 61 VERACMSHAIVLMKQALDGLPYITVFEDVLLGEGEKEFLADAWLQERFPDPTAFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
QY 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
QY 181 VDLMEFSPDRGMPVCOINPALCAQELHYAKFDONSALGSLIEHDLRLNKQQRDS 240
DB 181 VDLMEFSPDRGMPVCOINPALCAQELHYAKFDONSALGSLIEHDLRLNKQQRDS 240
QY 181 VDLMEFTYFDKGEKMPVQVSPALCTQELHYAKFLSQNSWLSGLDKD---REGRRRR 236
DB 181 VDLMEFTYFDKGEKMPVQVSPALCTQELHYAKFLSQNSWLSGLDKD---REGRRRR 236
QY 241 PANTFHKRLIRALTKISREKRRORREQ 269
DB 241 PANTFHKRLIRALTKISREKRRORREQ 269
QY 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 9
US-08-683-458-6
Sequence 6, Application US/08683458
Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-458-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACMSHAIVLMKQALDGLPYITVFEDVLLGEGEKEFLADAWLQERFPDPTAFYR 120
DB 61 VERACMSHAIVLMKQALDGLPYITVFEDVLLGEGEKEFLADAWLQERFPDPTAFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
QY 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPLLSESHWGTAGYIISRKARFFLDPRPALPREGGLHP 180
QY 181 VDLMEFSPDRGMPVCOINPALCAQELHYAKFDONSALGSLIEHDLRLNKQQRDS 240
DB 181 VDLMEFSPDRGMPVCOINPALCAQELHYAKFDONSALGSLIEHDLRLNKQQRDS 240
QY 181 VDLMEFTYFDKGEKMPVQVSPALCTQELHYAKFLSQNSWLSGLDKD---REGRRRR 236
DB 181 VDLMEFTYFDKGEKMPVQVSPALCTQELHYAKFLSQNSWLSGLDKD---REGRRRR 236
QY 241 PANTFHKRLIRALTKISREKRRORREQ 269
DB 241 PANTFHKRLIRALTKISREKRRORREQ 269
QY 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 10
US-08-878-360-6
Sequence 6, Application US/08878360
Patent No. 5945322
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,360
FILING DATE: 18-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/683,426
FILING DATE:

APPLICATION NUMBER: 08/312,387
 FILING DATE: September 26, 1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-0958
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 280 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-878-360-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
 Best Local Similarity 74.7%; Pred. No. 1.1e-109;
 Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MNNHVISLASAARRAHADTFGRGIPFPFDALMPSERLQAMAEIVPGISAHPIYLSG 60
 DB 1 MNNHVISLASAARRAHADTFGRGIPFPFDALMPSERLQAMAEIVPGISAHPIYLSG 60
 QY 61 VEKACFMSHAYLMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAMLOERPDPTAFIVR 120
 DB 61 VEKACFMSHAYLMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAMLOERPDPTAFIVR 120
 QY 121 LETMFHAYITSPSGVADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDPAALPEEGLHP 180
 DB 121 LETMFHAYITSPSGVADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDPAALPEEGLHP 180
 QY 181 VDIAMFSPFDEBEGMPVQALNPALCAQELHYAKFHDONSALGSLIEHRLNRKQORDS 240
 DB 181 VDIAMFSPFDEBEGMPVQALNPALCAQELHYAKFHDONSALGSLIEHRLNRKQORDS 240
 QY 241 PANTFKHRLIRALTISRERERKRRORQ 269
 DB 241 PANTFKHRLIRALTISRERERKRRORQ 269
 QY 237 RSLKVMFDLKRALGKGRKKRMRORQ 265
 DB 237 RSLKVMFDLKRALGKGRKKRMRORQ 265

RESULT 11

US-08-478-140B-6
 Sequence 6, Application US/08478140B
 Patent No. 6127153

GENERAL INFORMATION:
 APPLICANT: JOHNSON, KARL F.
 APPLICANT: ROTH, STEPHEN
 APPLICANT: BUCZALA, STEPHANIE L.
 TITLE OF INVENTION: METHOD OF TRANSPERRING AT LEAST TWO
 TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
 TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
 TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: IBM PC floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Releasee #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/478,140B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Laura A. Coruzzi
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7188-017
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 280 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-478-140B-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
 Best Local Similarity 74.7%; Pred. No. 1.1e-109;
 Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MNNHVISLASAARRAHADTFGRGIPFPFDALMPSERLQAMAEIVPGISAHPIYLSG 60
 DB 1 MNNHVISLASAARRAHADTFGRGIPFPFDALMPSERLQAMAEIVPGISAHPIYLSG 60
 QY 61 VEKACFMSHAYLMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAMLOERPDPTAFIVR 120
 DB 61 VEKACFMSHAYLMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAMLOERPDPTAFIVR 120
 QY 121 LETMFHAYITSPSGVADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDPAALPEEGLHP 180
 DB 121 LETMFHAYITSPSGVADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDPAALPEEGLHP 180
 QY 181 VDIAMFSPFDEBEGMPVQALNPALCAQELHYAKFHDONSALGSLIEHRLNRKQORDS 240
 DB 181 VDIAMFSPFDEBEGMPVQALNPALCAQELHYAKFHDONSALGSLIEHRLNRKQORDS 240
 QY 241 PANTFKHRLIRALTISRERERKRRORQ 269
 DB 241 PANTFKHRLIRALTISRERERKRRORQ 269
 QY 237 RSLKVMFDLKRALGKGRKKRMRORQ 265
 DB 237 RSLKVMFDLKRALGKGRKKRMRORQ 265

RESULT 12

US-09-333-412-6
 Sequence 6, Application US/09333412
 Patent No. 6342382

GENERAL INFORMATION:
 APPLICANT: Gotschlich, Emil C.
 TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Releasee #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/333,412
 FILING DATE: 15-Jun-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/312,387
 FILING DATE: July 7, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-095

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-333-412-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
QY 61 VERACPMASHAVIMKQALDEGLPYITVFEEDVLLGEGEKPFLABDAMLQERPDPTAIVR 120
DB 61 VERACPMASHAVIMKQALDEGLPYITVFEEDVLLGEGEKPFLABDAMLQERPDPTAIVR 120
QY 121 LETMFHNVLTSPSGVADYCGRAFPILSEHMGTAGYIISRKARFPFLDRPALPPEGILAP 180
DB 121 LETMFHNVLTSPSGVADYCGRAFPILSEHMGTAGYIISRKARFPFLDRPALPPEGILAP 180
QY 181 VDLMEFSDPFDRBGMVPCQALNPALCAQELHYAKFHQNSALGSLIEHDLNLRKQORSD 240
DB 181 VDLMEFSDPFDRBGMVPCQALNPALCAQELHYAKFHQNSALGSLIEHDLNLRKQORSD 240
QY 241 PANTFKHRLIRALTYISREKRRORREQ 269
DB 241 PANTFKHRLIRALTYISREKRRORREQ 269
QY 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 13

US-09-338-943-6
Sequence 6, Application US/09338943
Patent No. 6379933
GENERAL INFORMATION:
APPLICANT: JOHNSON, KARL F.
APPLICANT: BOTH, STEPHEN
APPLICANT: BUCZALA, STEPHANIE L.
TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
TITLE OF INVENTION: SACHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,943
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,140
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Cortuzzi
REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-338-943-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
QY 61 VERACPMASHAVIMKQALDEGLPYITVFEEDVLLGEGEKPFLABDAMLQERPDPTAIVR 120
DB 61 VERACPMASHAVIMKQALDEGLPYITVFEEDVLLGEGEKPFLABDAMLQERPDPTAIVR 120
QY 121 LETMFHNVLTSPSGVADYCGRAFPILSEHMGTAGYIISRKARFPFLDRPALPPEGILAP 180
DB 121 LETMFHNVLTSPSGVADYCGRAFPILSEHMGTAGYIISRKARFPFLDRPALPPEGILAP 180
QY 181 VDLMEFSDPFDRBGMVPCQALNPALCAQELHYAKFHQNSALGSLIEHDLNLRKQORSD 240
DB 181 VDLMEFSDPFDRBGMVPCQALNPALCAQELHYAKFHQNSALGSLIEHDLNLRKQORSD 240
QY 241 PANTFKHRLIRALTYISREKRRORREQ 269
DB 241 PANTFKHRLIRALTYISREKRRORREQ 269
QY 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 14

US-10-007-267A-6
Sequence 6, Application US/10007267A
Patent No. 6780624
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
FILE REFERENCE: 040853-01-5029-02
CURRENT APPLICATION NUMBER: US/10/007,267A
CURRENT FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US 09/333,412
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/878,360
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: US 08/683,426
PRIOR FILING DATE: 1996-07-18
PRIOR APPLICATION NUMBER: US 08/312,387
PRIOR FILING DATE: 1994-09-24
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 280
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-007-267A-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMAEVPGLSAHPYLSG 60

```
QY      61  VEKACFMSHAVLWKQALDEGLPYITVFPDDVLLGSGEKKFLAEDAMLOERFPDPTAFIVR 120
DB      61  VEKACFMSHAVLWKQALDEGLPYITVFPDDVLLGSGAEOFLAEDTWERFPDQSAFIVR 120
QY      121 LETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRPALPPEGILHP 180
DB      121 LETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRPALPPEGILHP 180
QY      181 VDIAMFSPDFDEGMPVQOLNPALCAOEIHYAKFHDONSALGSLIEHDLINRKQQRDS 240
DB      181 VDIAMFSPDFDEGMPVQOLNPALCAOEIHYAKFHDONSALGSLIEHDLINRKQQRDS 240
QY      241 PANTFKHRLIRALTKISREKERQRQRQ 269
DB      237 RSLKVMFDLKRALGKFGREKKRMRERQ 265
```

```
RESULT 15
US-09-540-236-2407
; Sequence 2407, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2407
; LENGTH: 255
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2407
```

```
Query Match      25.2%; Score 364; DB 2; Length 255;
Best Local Similarity 35.2%; Pred. No. 3.8e-34;
Matches 90; Conservative 49; Mismatches 103; Indels 14; Gaps 7;
```

```
QY      1  MGNHYISLASAEERRAHADTGRHGIPQFPDMLMPSERLEQAMAEIVPGLSAHPIYLSG 60
DB      3  IONFVIVTATKRRRHIMCFEGKQGIAPFPDAVPTDISRYAQKLSIPTINORLTDG 62
QY      61  VEKACFMSHAVLWKQALDEGLPYITVFPDDVLLGSGEKKFLAEDAMLOERFPDPTAFIV 119
DB      63  -EKACFMSHAVLWKQALDEGLPYITVFPDDVLLGSGEKKFLAEDAMLOERFPDPTAFIV 117
QY      120 RLETFMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRPALPPEGILHP 179
DB      118 RLETFMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRPALPPEGILHP 176
QY      180 PVDLMMFSPDFDEGMPVQOLNPALCAOEIHYAKFHDONSALGSLIEHDLINRKQQRDS 239
DB      177 PVDLMMFSPDFDEGMPVQOLNPALCAOEIHYAKFHDONSALGSLIEHDLINRKQQRDS 230
QY      240 SPANTFKHRLIRALTK 255
DB      231 TLAD-YGKKYRSIGK 245
```

```
Search completed: April 7, 2006, 15:31:27
Job time : 49 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:41:34 ; Search time 168 Seconds
(Without alignments)
683.947 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MGNHVISLSAARRAHAD.....ISRERKRQRQRFIVPQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
Published Applications AA Main:
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1446	100.0	275	3	US-09-211-691-2
2	1446	100.0	275	4	US-10-317-773-2
3	1446	100.0	275	4	US-10-317-428-2
4	1288	89.1	279	4	US-10-007-267-8
5	1288	89.1	279	4	US-10-654-528-8
6	1007	69.6	280	4	US-10-007-267-6
7	1007	69.6	280	4	US-10-096-129-6
8	1007	69.6	280	4	US-10-654-528-6
9	933	64.5	280	5	US-10-472-260-176
10	130.5	9.0	622	4	US-10-398-037-1
11	128.5	8.9	332	4	US-10-335-977-8513
12	122.5	8.5	1128	5	US-11-097-143-15177
13	121	8.4	738	5	US-10-723-860-1579
14	118.5	8.2	423	4	US-10-264-237-1829
15	118.5	8.2	560	5	US-10-723-860-4111
16	118.5	8.2	579	4	US-10-062-548-61
17	118.5	8.2	579	5	US-10-918-446-61
18	118.5	8.2	579	6	US-11-002-755-61
19	118.5	8.2	579	6	US-11-002-756-61
20	118.5	8.2	636	4	US-10-264-049-3148
21	110	7.6	201	4	US-10-335-977-8512
22	110	7.6	739	4	US-10-243-552-866
23	110	7.6	739	4	US-10-276-774-1787
24	104	7.2	762	6	US-11-035-599-28
25	100.5	7.0	168	5	US-10-450-763-52449
26	98.5	6.8	273	4	US-10-218-381-1
27	98	6.8	481	4	US-10-475-587-1

28	97.5	6.7	133	4	US-10-335-977-8517	Sequence 8517, Ap
29	97.5	6.7	144	4	US-10-335-977-8518	Sequence 8518, Ap
30	93	6.4	109	4	US-10-251-0858-241	Sequence 241, App
31	93	6.4	109	4	US-10-737-252-241	Sequence 241, App
32	93	6.4	113	5	US-10-700-632-78	Sequence 78, App1
33	93	6.4	150	3	US-09-226-157-2	Sequence 2, App1
34	93	6.4	237	4	US-10-336-210-2	Sequence 2, App1
35	93	6.4	237	4	US-10-336-210-3	Sequence 3, App1
36	93	6.4	211	3	US-09-326-157-4	Sequence 4, App1
37	93	6.4	271	4	US-10-336-210-6	Sequence 6, App1
38	93	6.4	272	4	US-10-336-210-7	Sequence 7, App1
39	92	6.4	244	3	US-09-940-391-1	Sequence 1, App1
40	92	6.4	244	4	US-10-336-210-8	Sequence 8, App1
41	92	6.4	483	5	US-10-858-367-17	Sequence 17, App1
42	91	6.3	133	4	US-10-355-478-48	Sequence 48, App1
43	91	6.3	242	4	US-10-336-210-9	Sequence 9, App1
44	90	6.2	113	4	US-10-251-0858-235	Sequence 235, App
45	90	6.2	113	4	US-10-737-252-235	Sequence 235, App

ALIGNMENTS

RESULT 1	US-09-211-691-2	
Sequence 2, Application US/09211691		
Patent No. US20020034805A1		
GENERAL INFORMATION:		
APPLICANT: Gilbert, Michel		
APPLICANT: Young, N. Martin		
APPLICANT: Makarchuk, Warren W.		
TITLE OF INVENTION: National Research Council of Canada		
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of		
FILE REFERENCE: 019957-012910US		
CURRENT APPLICATION NUMBER: US/09/211,691		
CURRENT FILING DATE: 1998-12-14		
PRIOR APPLICATION NUMBER: US 60/069,443		
PRIOR FILING DATE: 1997-12-15		
NUMBER OF SEQ ID NOS: 18		
SOFTWARE: PatentIn Ver. 2.1		
SEQ ID NO 2		
LENGTH: 275		
TYPE: PRT		
ORGANISM: Neisseria meningitidis		
US-09-211-691-2		
Query Match	100.0%	Score 1446; DB 3; Length 275;
Best Local Similarity	100.0%	Pred. No. 3.1e-141;
Matches 275; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	MGNHVISLSAARRAHADTFGRHGIPIPOFPDAMPSSRLROAMAEVPGLSAPPYLSG 60
DB	1	MGNHVISLSAARRAHADTFGRHGIPIPOFPDAMPSSRLROAMAEVPGLSAPPYLSG 60
QY	61	VEKAFPMGSHAVIMKQALBEGLPYITVFBDVLLSGEEXFPAEDWMLQRPDPDTAIVR 120
DB	61	VEKAFPMGSHAVIMKQALBEGLPYITVFBDVLLSGEEXFPAEDWMLQRPDPDTAIVR 120
QY	121	LETMFMAHVLTSQGVADYCGRAFPILLESBHWTAGYIISRKAMRFFLDRFALPEGLHP 180
DB	121	LETMFMAHVLTSQGVADYCGRAFPILLESBHWTAGYIISRKAMRFFLDRFALPEGLHP 180
QY	181	VDLMMFSDFDEGMPVQOLNPAACAELHYAKFDONSALGSLTEHRLINRKOQRDS 240
DB	181	VDLMMFSDFDEGMPVQOLNPAACAELHYAKFDONSALGSLTEHRLINRKOQRDS 240
QY	241	PANTFKHLIRALTITISREKRQRQRFIVPQ 275
DB	241	PANTFKHLIRALTITISREKRQRQRFIVPQ 275

RESULT 2

US-10-317-773-2
; Sequence 2, Application US/10317773
; Publication No. US20030180928A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNac 4' Epimerase and a
; FILE REFERENCE: GalNac Transferase
; FILE REFERENCE: 019633-00081205
; CURRENT APPLICATION NUMBER: US/10/317,773
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-317-773-2

Query Match 100.0%; Score 1446; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.1e-141;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNHVISLASAARRAHADTFGRGIPFOFPDMLMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MNNHVISLASAARRAHADTFGRGIPFOFPDMLMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAYLMKQALDEGLPYITVFEDVDVLGSEGEKFLAEDAWLOERPDPTAFIVR 120
DB 61 VEKACFMSHAYLMKQALDEGLPYITVFEDVDVLGSEGEKFLAEDAWLOERPDPTAFIVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSEHMGTAIISRKAMRFFLDLRFPAALPPEGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSEHMGTAIISRKAMRFFLDLRFPAALPPEGGLHP 180
QY 181 VDLAMFSDPFDREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLAMFSDPFDREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFKHRLIRALTIXISREKRRQRREQFIVPFQ 275
DB 241 PANTFKHRLIRALTIXISREKRRQRREQFIVPFQ 275

RESULT 3
US-10-317-428-2
; Sequence 2, Application US/10317428
; Publication No. US2003018641A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
; FILE REFERENCE: 019633-00081105
; CURRENT APPLICATION NUMBER: US/10/317,428
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Neisseria meningitidis

US-10-317-428-2

Query Match 100.0%; Score 1446; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.1e-141;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNHVISLASAARRAHADTFGRGIPFOFPDMLMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MNNHVISLASAARRAHADTFGRGIPFOFPDMLMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAYLMKQALDEGLPYITVFEDVDVLGSEGEKFLAEDAWLOERPDPTAFIVR 120
DB 61 VEKACFMSHAYLMKQALDEGLPYITVFEDVDVLGSEGEKFLAEDAWLOERPDPTAFIVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSEHMGTAIISRKAMRFFLDLRFPAALPPEGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSEHMGTAIISRKAMRFFLDLRFPAALPPEGGLHP 180
QY 181 VDLAMFSDPFDREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLAMFSDPFDREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFKHRLIRALTIXISREKRRQRREQFIVPFQ 275
DB 241 PANTFKHRLIRALTIXISREKRRQRREQFIVPFQ 275

RESULT 4
US-10-007-267-8
; Sequence 8, Application US/10007267
; Publication No. US20020127682A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Klauder & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,267
; FILING DATE: 03-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELE: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-007-267-8

Query Match 89.1%; Score 1288; DB 4; Length 279;
Best Local Similarity 90.0%; Pred. No. 7,8e-125;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 M0NHV1SLASAERRAH1ADTFGRHGI PFOFPDAMPSERLERQAMAEIVPGISAHPIYLSG 60
DB 1 M0NHV1SLASAERRAH1ATFGSRGIPFOFPDAMPSERLERQAMAEIVPGISAHPIYLSG 60
QY 61 VERACFMSHAIVLMKQALDEGLPYITVPEDVLLGSEBEKFLADAMLOERPDPTAFIYR 120
DB 61 VERACFMSHAIVLMKQALDEGLPYITVPEDVLLGSEBEKFLADAMLOERPDPTAFIYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSESHMGTAAGIISRKAMRFFLDRFALPPEGILHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSESHMGTAAGIISRKAMRFFLDRFALPPEGILHP 180
QY 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRHLIRALTITKISREKRRORRQ----FIVPFQ 275
DB 241 PANTFGRHLIRALTITKISREKRRORRQELIGKITIVPFQ 279

RESULT 5

US-10-654-528-8
Sequence 8, Application US/10654528
Publication No. US20040043464A1
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
FILE REFERENCE: 040853-01-5029-03
CURRENT APPLICATION NUMBER: US/10/654,528
PRIOR FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US 10/007,267
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US 09/333,412
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/878,360
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: US 08/583,426
PRIOR FILING DATE: 1996-07-18
PRIOR APPLICATION NUMBER: US 08/312,387
PRIOR FILING DATE: 1994-09-24
SOFTWARE: Patentin version 3.2
SEQ ID NO 8
LENGTH: 279
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-654-528-8

Query Match 89.1%; Score 1288; DB 4; Length 279;
Best Local Similarity 90.0%; Pred. No. 7,8e-125;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 M0NHV1SLASAERRAH1ADTFGRHGI PFOFPDAMPSERLERQAMAEIVPGISAHPIYLSG 60
DB 1 M0NHV1SLASAERRAH1ATFGSRGIPFOFPDAMPSERLERQAMAEIVPGISAHPIYLSG 60
QY 61 VERACFMSHAIVLMKQALDEGLPYITVPEDVLLGSEBEKFLADAMLOERPDPTAFIYR 120
DB 61 VERACFMSHAIVLMKQALDEGLPYITVPEDVLLGSEBEKFLADAMLOERPDPTAFIYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSESHMGTAAGIISRKAMRFFLDRFALPPEGILHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSESHMGTAAGIISRKAMRFFLDRFALPPEGILHP 180
QY 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240

DB 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRHLIRALTITKISREKRRORRQ----FIVPFQ 275
DB 241 PANTFGRHLIRALTITKISREKRRORRQELIGKITIVPFQ 279

RESULT 6

US-10-007-267-6
Sequence 6, Application US/10007267
Publication No. US20020127682A1
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-007-267-6

Query Match 69.6%; Score 1007; DB 4; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-95;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 M0NHV1SLASAERRAH1ADTFGRHGI PFOFPDAMPSERLERQAMAEIVPGISAHPIYLSG 60
DB 1 M0NHV1SLASAERRAH1ATFGSRGIPFOFPDAMPSERLERQAMAEIVPGISAHPIYLSG 60
QY 61 VERACFMSHAIVLMKQALDEGLPYITVPEDVLLGSEBEKFLADAMLOERPDPTAFIYR 120
DB 61 VERACFMSHAIVLMKQALDEGLPYITVPEDVLLGSEBEKFLADAMLOERPDPTAFIYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSESHMGTAAGIISRKAMRFFLDRFALPPEGILHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSESHMGTAAGIISRKAMRFFLDRFALPPEGILHP 180
QY 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDPFDRGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240

Qy 241 PANTFKHLIRALTYISRRERKRRREQ 269
Db 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 7

US-10-096-129-6
; Sequence 6, Application US/10096129
; Publication No. US20030207406A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: ROTH, STEPHEN
; APPLICANT: BUCCALA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO SACCHARIDE UNITS WITH A
; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
; FILE REFERENCE: 040853-01-5021-02
; CURRENT APPLICATION NUMBER: US/10/096,129
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/338,943
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 08/478,140
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 280
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-096-129-6

Query Match 69.6%; Score 1007; DB 4; Length 280;

Best Local Similarity 74.7%; Pred. No. 1.1e-95; Mismatches 44; Indels 4; Gaps 1;

Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

Qy 1 MONHVISLSAARAHADTFGRHGIPPOFFDALMPSERLQAMAEVPGLSAHPTYLSG 60
Db 1 MONHVISLSAARAHADTFGRHGIPPOFFDALMPSERLQAMAEVPGLSAHPTYLSG 60
Qy 61 VEKACFMSHAYVLMKQALDEGLPYITVFEEDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
Db 61 VEKACFMSHAYVLMKQALDEGLPYITVFEEDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
Qy 121 LETMFHVLITSPSGVADYCGRAFPILSEHMGCTAGYIISRKAMRFFLDPAALPPEGGLHP 180
Db 121 LETMFHVLITSPSGVADYCGRAFPILSEHMGCTAGYIISRKAMRFFLDPAALPPEGGLHP 180
Qy 181 VDLMPFSDPFDREGMPCVQNLPAALCAQELHYAKFHDONSALGSLIENDRLNRKQORRDS 240
Db 181 VDLMPFSDPFDREGMPCVQNLPAALCAQELHYAKFHDONSALGSLIENDRLNRKQORRDS 240
Qy 241 PANTFKHLIRALTYISRRERKRRREQ 269
Db 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 8

US-10-654-528-6
; Sequence 6, Application US/10654528
; Publication No. US20040043464A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 6
; LENGTH: 280
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-654-528-6

Query Match 69.6%; Score 1007; DB 4; Length 280;

Best Local Similarity 74.7%; Pred. No. 1.1e-95; Mismatches 44; Indels 4; Gaps 1;

Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

Qy 1 MONHVISLSAARAHADTFGRHGIPPOFFDALMPSERLQAMAEVPGLSAHPTYLSG 60
Db 1 MONHVISLSAARAHADTFGRHGIPPOFFDALMPSERLQAMAEVPGLSAHPTYLSG 60
Qy 61 VEKACFMSHAYVLMKQALDEGLPYITVFEEDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
Db 61 VEKACFMSHAYVLMKQALDEGLPYITVFEEDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
Qy 121 LETMFHVLITSPSGVADYCGRAFPILSEHMGCTAGYIISRKAMRFFLDPAALPPEGGLHP 180
Db 121 LETMFHVLITSPSGVADYCGRAFPILSEHMGCTAGYIISRKAMRFFLDPAALPPEGGLHP 180
Qy 181 VDLMPFSDPFDREGMPCVQNLPAALCAQELHYAKFHDONSALGSLIENDRLNRKQORRDS 240
Db 181 VDLMPFSDPFDREGMPCVQNLPAALCAQELHYAKFHDONSALGSLIENDRLNRKQORRDS 240
Qy 241 PANTFKHLIRALTYISRRERKRRREQ 269
Db 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 9

US-10-472-260-176
; Sequence 176, Application US/10472260
; Publication No. US20040265328A1
; GENERAL INFORMATION:
; APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY
; APPLICANT: IMPERIAL COLLEGE INNOVATIONS LIMITED
; TITLE OF INVENTION: IMMUNOGENIC COMMENSAL NEISSERIA SEQUENCES
; FILE REFERENCE: GMS/DJC/23480
; CURRENT APPLICATION NUMBER: US/10/472,260
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 176
; LENGTH: 280
; TYPE: PRF
; ORGANISM: Neisseria meningitidis (group B)
US-10-472-260-176

Query Match 64.5%; Score 933; DB 5; Length 280;

Best Local Similarity 69.5%; Pred. No. 5.3e-88; Mismatches 54; Indels 4; Gaps 1;

Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

Qy 1 MONHVISLSAARAHADTFGRHGIPPOFFDALMPSERLQAMAEVPGLSAHPTYLSG 60
Db 1 MONHVISLSAARAHADTFGRHGIPPOFFDALMPSERLQAMAEVPGLSAHPTYLSG 60
Qy 61 VEKACFMSHAYVLMKQALDEGLPYITVFEEDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
Db 61 VEKACFMSHAYVLMKQALDEGLPYITVFEEDVVLGGESEKFLAEDAMLOERPDPTAFIVR 120
Qy 121 LETMFHVLITSPSGVADYCGRAFPILSEHMGCTAGYIISRKAMRFFLDPAALPPEGGLHP 180
Db 121 LETMFHVLITSPSGVADYCGRAFPILSEHMGCTAGYIISRKAMRFFLDPAALPPEGGLHP 180
Qy 181 VDLMPFSDPFDREGMPCVQNLPAALCAQELHYAKFHDONSALGSLIENDRLNRKQORRDS 240
Db 181 VDLMPFSDPFDREGMPCVQNLPAALCAQELHYAKFHDONSALGSLIENDRLNRKQORRDS 240

Db 181 VDMWMTYFDFKGMFVYQVNPALCTOELHYAKFLSKNSWLGSDLEKD----RQERRHR 236
QY 241 PAMTFKRLIRALTXTSREBRKRRQREQ 269
Db 237 RSLKVMFDLKRALGKGRKKRKKRQREQ 265

RESULT 10

US-10-398-037-1
Sequence 1, Application US/10398037
Publication No. US20040138414A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS; YUE, Henry;
TANG, Y. Tom; NGUYEN, Daniel B.;
VANO, Monique G.; XU, Yuming;
APPLICANT: TRIBOULET, Catherine M.; SANJANNALA, Madhusudan M.;
CHAMLA, Narinder K.; BAUGH, Mariah R.;
APPLICANT: SAPPERSTEIN, Stephanie K.; LAL, Preeti G.;
THORNTON, Michael B.; GANDHI, Ameeta R.;
APPLICANT: RAMKUMAR, Jayalaxmi; ELIOTT, Vicki S.;
ARVIZU, Chandra S.; THANGAVELU, Kavitha;
APPLICANT: GIEZEN, Kimberly J.; DING, Li;
APPLICANT: AU-YOUNG, Janice K.; TRAN, Bao;
APPLICANT: POLICKY, Jennifer L.; LEE, Sally;
APPLICANT: LU, Dying Aina M.; BURFORD, Neil;
APPLICANT: WARREN, Bridget A.; GURURAJAN, Rajagopal;
APPLICANT: DUGGAN, Brendan M.; HONCHELL, Cynthia D.;
APPLICANT: HAFALIA, April J.A.
TITLE OF INVENTION: SECRETED PROTEINS
FILE REFERENCE: PI-0240 USN
CURRENT APPLICATION NUMBER: US/10/398,037
PRIOR FILING DATE: 2003-03-28
PRIOR APPLICATION NUMBER: PCT/US01/30042
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/242,218
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/236,869
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/239,812
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: US 60/240,108
PRIOR FILING DATE: 2000-10-12
PRIOR APPLICATION NUMBER: US 60/241,282
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 136
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 622
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 3211795CD1
US-10-398-037-1

Query Match 9.0%; Score 130.5; DB 4; Length 622;
Best Local Similarity 25.1%; Pred. No. 0.00029;
Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;
QY 5 VLSIASAABERRAHITDFTGNGIIPQFPDNL-----WPSRLBOAMAEIVPGLSAHFY--- 57
Db 345 MTLRRRQDRREKMLBALQOBIJECRLVEAVDGGKAMTSGVBLGIGMLFGY-RDPYHGR 403
QY 58 -LSGVKACPMASHAVLMKQALDEGLPYITVFEDV-----LIGRGEKEFLAED 104
Db 404 PLTKGELGCLSHYNNIMKAVVVDGLQKSLVFEDDLRPEIFFKRRLNMIMKWDVE----- 456
QY 105 AMIQERPDPTAFIVLETMFVGLTSPSGVADYCGRAP-----LLESEH-WCTAGYIIS 159
Db 457 ---REGIDMDLIVYGGKRMQVH---FE-----KAVPRVNNLVADYSYVTLAYVIS 502
QY 160 RKAMRPFLLDRFALPPEGILHVDLMMFSDPDRGMPVQGLNPAICQGEIH 210

Db 503 LQARLTL---AAEFISKQLPVDEFL-PVWFDRH--PVSEYKAHFSIRNLH 547

RESULT 11

US-10-335-977-8513
Sequence 8513, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHYE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-Dec-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 8513:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...332
SEQUENCE DESCRIPTION: SEQ ID NO: 8513:
US-10-335-977-8513

Query Match 8.9%; Score 128.5; DB 4; Length 332;
Best Local Similarity 23.5%; Pred. No. 0.0002; Indels 103; Gaps 17;
Matches 74; Conservative 39; Mismatches 99;
QY 29 PQFPDMLPSEK-LBOAMAEIVGLSA-----HPY-----LSGVKACPMASHAVLMKQAL 77
Db 35 PQFPDMLSPKHQDFELQLOELVNAQSLQSDWYHSVVGAGTLIPELGCTLSHLMKECV 94
QY 78 DRGLPYITVFEDVVLIGBGEKFL-AEDAMIQERPDPTAFIVRL-----ETMFMEY 128
Db 95 KLDPQ-VVILBDVTL---ESHFMQLBDCIKSPD-----FVRLVGYCYWYHETKF-HV 144
QY 129 LNSPSGVA---DYCRARFPLES----- 148
Db 145 L-PKEFVPPDPDHSKNNPILKPKFPDVSRLNLSTHKVYIYTLKKIQNSYATHEK 202
QY 149 -----EHW-----GTAGYIISRKAMRFLDRFALPPEGILHVDLMMFSDPDRBGM 195
Db 203 EAPFLEHFLTSSVYVASTAGYTLTPKGAKTFEATESF--KILFVDMFMDSAHY----- 256

QY 196 PVOCLNPAALCAQSLAHAKFHQDSALGSLIHEDRLNLRMOGRSDSPATFGHRLIALTK 255
Db 257 DVANLTVPCEPVSLS-----SEHSLDSLTITQRPQKSLKSYLPAPQSKTFN-----LFY 304

QY 256 ISRRERERRRRRRP 270
Db 305 YSLNAKGRRLNAFOQY 319

RESULT 12
US-11-097

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	8.5%;	Score 122.5;	DB 6;	Length 1128;
Best Local Similarity	25.7%;	Pred. No. 0.0043;		
Matches	56;	Conservative	26;	Mismatches 73;
				Indels 63;
				Gaps 10

```

QY      5  VISLISAERKRIHIDTGRHSIIPQPFQAL-----MPSRRLEQAAELVPGUS---APHY  57
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
Db      653  MINLKRREERREKRMRLFDEIGIEAHNPAYDGKELSTRLLRMGVRLPGIPEDDYYHRA  912
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
QY      58  LSGVEKACFMSHVALMKQALDEGLPYITVFEEDVLLIGBEEKFLAEDAWLQERFDPDTAF  117
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
Db      913  MTMGELGGCFLSHYNIWMVVVRKQLMEVILBEDI-----RPEE-----  950
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
QY      118  IVRLETTMAHVLITSSGVAD---YCGRAFLPILBSE-----HWGTAGYII  158
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
Db      951  --YPRQNAVRILNQARNAQYDLIYFGRRKLKEBSEPAVENADNLVHAGYSW-TLGYVI  1007
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
QY      159  SRKARFFLDREFAALPBGHLPVD---LMMKSDPFDR  192
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||
Db      1008  S---LQGALKLIAAKPLDGLIPVDEFLPLM-----FDR  1037
      ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||

```

RESULT 13
US-10-723-860-1579
; Sequence 1579, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:

```

1  APPLICANT: Aziz, Natasha
2  APPLICANT: Ginsburg, Wendy M.
3  APPLICANT: Zlotnick, Albert
4  TITLE OR INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
5  TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
6  FILE REFERENCE: 05882.0193.NPUS01
7  CURRENT APPLICATION NUMBER: US/10/723,860
8  CURRENT FILING DATE: 2003-11-26
9  PRIOR APPLICATION NUMBER: 60/429,739
10 PRIOR FILING DATE: 2002-11-26
11 NUMBER OF SEQ ID NOS: 8393
12 SOFTWARE: PatentIn version 3.2
13 SEQ ID NO 1579
14 LENGTH: 738
15 TYPE: PRT
16 ORGANISM: Homo sapiens
17 US-10-723-860-1579

```

Query Match	8.4%;	Score 121;	DB 5;	Length 738;
Best Local Similarity	23.1%;	Pred. No. 0.0035;		
Matches	53;	Conservative	42;	Mismatches 100;
				Indels 34;
				Gaps 9

QY 5 VISLISAERPAHIAADTGRGIGIPQPFADAL-----MPSERIEQAAAEIIVPG-----LSAP 56
 458 MINIKRRDRDRMRITLYEQEIEVKIYEAADGKALNTSOLKALINIMLPGRDRPYSSRP 517
 QY YLSGVEKACPMASHAVLKQALDEGLPYITTFEDDYLLABGEKEFLAE--DAWLQERFPDP 114
 518 LTRG-EIGCFPSHYSVMKEVIDRELEKTLVIEDVRFPHQFKKMLKMDNIDQADJWE 576
 QY TAFIVLETMMAHVLTPSGVADYVGRAFPLLIESH-MGTAGYIISRKAMRFFLDRLFALL 173
 577 LIYIGRKMQVKEPEKAVPNVAN-----LVEADYSWTLLGYVLSLEGAQKLV--GAN 626
 QY PREGHAPDULM-----PSDFPRDREGMPVQCLNPAACAGELHY 211
 Db PFGRKLVPDEFLPVWYNNGPVAAYEYKESRDLKAFSAEP-LLIYPTHY 674

RESULT 14
US-10-264-237-1829
; Sequence 1829, Application US/10264237
; Publication No. US20040009491A1
SUMMARY INFORMATION

```

? APPLICANT: Birex et al.
? TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
? FILE REFERENCE: P4133P1
? CURRENT APPLICATION NUMBER: US/10/264,237
? CURRENT FILING DATE: 2002-10-04
? PRIOR APPLICATION NUMBER: PCT/US01/16450
? PRIOR FILING DATE: 2001-05-18
? PRIOR APPLICATION NUMBER: US 60/205,515
? PRIOR FILING DATE: 2000-05-19
? NUMBER OF SEQ ID NOS: 2876
? SOFTWARE: PatentIn Ver. 3.1
? SEQ ID NO 1829
? LENGTH: 423
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-264-237-1829

```

Query Match	8.2%;	Score 118.5;	DB 4;	Length 423;
Best Local Similarity	26.2%;	Pred. No. 0.003;		
Matches 64;	Conservative 28;	Mismatches 89;	Indels 63;	Gaps 13

```

OY      5  VISLISAERPAHILDTGRGCIPOFPEDA-----LMPSERLQAMAEVPGLSAHY 57
Db      155 VISLARDRDRERMLASLMEWISGRVYDADGMMNLNSALINLGVLDLPQVQ-DYSGR 213
OY      58 -LSGVKRCFMSHVAVLMKQALDEGLPYTTVEEDDYVLGBGEKFLAADMALQERFPDPA 116
Db      214 TLTGKEVCCFLSHSIMEEVARGLARLAVPEDDV-----RFSN-----253

```

QY 117 FIVRLTETMFHVLITSP-SGVADYCGRAFPILISE-----HMGTAGYIISR 160
 DB 254 FRGRLERLMEVDYAEKLSMDLIYIGRKQVPEKETAVEGLPGLVVAGYSYV-TLAYAALRL 312
 QY 161 KAMRFFLDRAALPPEGHLPVD---LMF-----SDFFDRBGMPCQLNPALCAOE 208
 DB 313 AGARKLL---ASQPLRRMLPVDDEFLPIMFDQHPNEQYKAHFWPRD-LVAFSAQPLLA--P 367
 QY 209 LHYA 212
 DB 368 THYA 371

RESULT 15
 US-10-723-860-4111
 ; Sequence 4111, Application US/10723860
 ; Publication No. US20040253606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsburg, Wendy M.
 ; APPLICANT: Zlotnick, Albert
 ; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
 ; FILE REFERENCE: 05862.0193.NPUS01
 ; CURRENT APPLICATION NUMBER: US/10/723,860
 ; PRIOR FILING DATE: 2003-11-26
 ; PRIOR APPLICATION NUMBER: 60/429,739
 ; PRIOR FILING DATE: 2002-11-26
 ; NUMBER OF SEQ ID NOS: 8393
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4111
 ; LENGTH: 560
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-723-860-4111

Query Match 8.2%; Score 118.5; DB 5; Length 560;
 Best Local Similarity 26.2%; Pred. No. 0.0044;
 Matches 64; Conservative 28; Mismatches 89; Indels 63; Gaps 13;

QY 5 VISLSAARRAHIADTGRHGIPFOFPDA---LMPSERLQAMALVPGLSAHFY--- 57
 DB 323 VISLARPPRRRRBSMLASLWEMEISGRVVDVAVDGMMLNSSAIRMLGVDLPGYQ-DPYSGR 381
 QY 58 -LSGVKACGFMGSHAVLMKQALDEGLPYITVFPDDVULGSGEKEKFLAEDAMLOERPPDPA 116
 DB 382 TIKGIVGCFLSHSYSIWEEVARGLARVLVFPDDV-----RFRSN-- 421
 QY 117 FIVRLTETMFHVLITSP-SGVADYCGRAFPILISE-----HMGTAGYIISR 160
 DB 422 FRGRLERLMEVDYAEKLSMDLIYIGRKQVPEKETAVEGLPGLVVAGYSYV-TLAYAALRL 480
 QY 161 KAMRFFLDRAALPPEGHLPVD---LMF-----SDFFDRBGMPCQLNPALCAOE 208
 DB 481 AGARKLL---ASQPLRRMLPVDDEFLPIMFDQHPNEQYKAHFWPRD-LVAFSAQPLLA--P 535
 QY 209 LHYA 212
 DB 536 THYA 539

Search completed: April 7, 2006, 15:45:08
 Job time : 170 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:42:34 ; Search time 25 Seconds

(without alignments)
343.112 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHYSLASAAERRAHAD.....ISRRERKRQRREQIVPRQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSS/ptodata/1/pubppa/US08 NEW PUB.pep:*
2: /SIDSS/ptodata/1/pubppa/US06 NEW PUB.pep:*
3: /SIDSS/ptodata/1/pubppa/US07 NEW PUB.pep:*
4: /SIDSS/ptodata/1/pubppa/PCT_NEW PUB.pep:*
5: /SIDSS/ptodata/1/pubppa/US09 NEW PUB.pep:*
6: /SIDSS/ptodata/1/pubppa/US10 NEW PUB.pep:*
7: /SIDSS/ptodata/1/pubppa/US11 NEW PUB.pep:*
8: /SIDSS/ptodata/1/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1288	89.1	279	7	US-11-102-497-8
2	1281	88.6	279	6	US-10-467-657-1310
3	1007	69.6	280	7	US-11-102-497-6
4	996	68.9	280	6	US-10-467-657-1336
5	118.5	6.8	327	6	US-10-821-234-884
6	98	6.1	481	7	US-11-227-177-1
7	88.5	6.1	651	7	US-11-087-099-2611
8	87.5	6.1	238	7	US-11-239-510-21
9	87.5	6.1	290	7	US-11-032-773-957
10	87.5	6.1	319	7	US-11-032-773-955
11	84	5.8	395	7	US-11-084-220-2
12	83.5	5.8	440	7	US-11-239-510-11
13	83.5	5.8	483	7	US-11-239-510-19
14	83.5	5.8	353	6	US-10-506-454-1626
15	82	5.7	120	6	US-10-932-334-77
16	81.5	5.6	416	6	US-10-501-035-279
17	80.5	5.6	571	7	US-11-087-099-7090
18	79.5	5.5	119	7	US-11-226-886-20
19	79.5	5.5	119	7	US-11-159-046-4
20	79.5	5.5	248	6	US-10-512-184-32
21	79.5	5.5	317	6	US-10-512-184-69
22	79	5.5	425	7	US-11-096-568A-20512
23	79	5.5	645	7	US-11-096-568A-19117
24	79	5.5	648	7	US-11-096-568A-19116
25	79	5.5	711	7	US-11-096-568A-19115

26	78	5.4	331	7	US-11-087-099-3050	Sequence 3050, Ap
27	77.5	5.4	968	6	US-10-501-035-219	Sequence 219, App
28	77	5.3	363	7	US-11-121-731A-1	Sequence 1, Appl
29	77	5.3	539	6	US-10-793-626-340	Sequence 340, App
30	77	5.3	805	6	US-10-927-641-77	Sequence 77, Appl
31	76	5.3	118	6	US-10-932-334-75	Sequence 75, Appl
32	75.5	5.2	434	7	US-11-087-099-11503	Sequence 11503, A
33	75.5	5.2	687	6	US-10-485-517-274	Sequence 274, App
34	75	5.2	252	7	US-11-072-512-2120	Sequence 2120, App
35	75	5.2	268	7	US-11-096-568A-2036	Sequence 2036, Ap
36	75	5.2	805	6	US-10-485-517-188	Sequence 198, App
37	75	5.2	842	7	US-11-096-568A-34417	Sequence 34417, A
38	75	5.2	869	7	US-11-096-568A-34416	Sequence 34416, A
39	75	5.2	970	7	US-10-932-334-79	Sequence 34415, A
40	74	5.1	120	6	US-11-096-568A-14763	Sequence 79, Appl
41	74	5.1	303	7	US-11-096-568A-14762	Sequence 14763, A
42	74	5.1	395	7	US-11-096-568A-14762	Sequence 14762, A
43	74	5.1	411	7	US-11-096-568A-14761	Sequence 14761, A
44	73.5	5.1	302	7	US-11-165-226-127	Sequence 127, App
45	73.5	5.1	436	7	US-11-096-568A-21599	Sequence 21599, A

ALIGNMENTS

RESULT 1									
US-11-102-497-8									
; Sequence 8, Application US/11102497									
; Publication No. US20050271690A1									
; GENERAL INFORMATION:									
; APPLICANT: Gotschlich, Emil C.									
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND									
; FILE REFERENCE: 040853-01-5029-03									
; CURRENT FILING DATE: 2005-04-08									
; PRIOR APPLICATION NUMBER: US/10/654,528									
; PRIOR FILING DATE: 2003-09-02									
; PRIOR APPLICATION NUMBER: US 10/007,267									
; PRIOR FILING DATE: 2001-12-03									
; PRIOR APPLICATION NUMBER: US 09/333,412									
; PRIOR FILING DATE: 1999-06-15									
; PRIOR APPLICATION NUMBER: US 08/878,360									
; PRIOR FILING DATE: 1997-06-18									
; PRIOR APPLICATION NUMBER: US 08/683,426									
; PRIOR FILING DATE: 1996-07-18									
; PRIOR APPLICATION NUMBER: US 08/312,387									
; PRIOR FILING DATE: 1994-09-24									
; NUMBER OF SEQ ID NOS: 13									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 8									
; LENGTH: 279									
; TYPE: PRT									
; ORGANISM: Neisseria gonorrhoeae									
US-11-102-497-8									
Query Match 89.1%; Score 1288; DB 7; Length 279;									
Best Local Similarity 90.0%; Pred. No. 5.4e-17; Indels 4; Gaps 1;									
Matches 251; Conservative 7; Mismatches 12;									
QY	1	MGNHYSLASAAERRAHADTFGRHGRTPFQFDALMPSERLERAAKLVPGLSAHPTVSG	60						
DB	1	MGNHYSLASAAERRAHADTFGRHGRTPFQFDALMPSERLERAAKLVPGLSAHPTVSG	60						
QY	61	VEKACFMSHAVLWQALDEGLPYITVPEBDVLGSGEKEFLAEDAMLOERFPDPTAFIVR	120						
DB	61	VEKACFMSHAVLWQALDEGLPYITVPEBDVLGSGEKEFLAEDAMLOERFPDPTAFIVR	120						
QY	121	LETMTMTLTPSGVADYVCGRAFLPESHEHGTGYTISRKAMPFLDRPALPPEGCHLP	180						
DB	121	LETMTMTLTPSGVADYVCGRAFLPESHEHGTGYTISRKAMPFLDRPALPPEGCHLP	180						
QY	181	VDLMMFSDFFREGMPVCQALPALCAQELHYAKFHQDSALGSLIEHRLNLRKQQRDS	240						

```

Db      181 VDLMMFGNDDREGMPVCOINPALCAQELHYAKFHDONSALGLEHDLRLNRKQWRDS 240
Qy      241 PANTFKHRLIRALTITKISREERKRRQREQ----FTVFPQ 275
        241 PANTFKHRLIRALTITKIGREERKRRQREQGLIKIIVFPQ 279

```

```

RESULT 2
US-10-467-657-1310
; Sequence 1310, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1310
; LENGTH: 279
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1310

```

```

Query Match      88.6%; Score 1281; DB 6; Length 279;
Best Local Similarity 90.0%; Pred. No. 2.7e-119;
Matches 251; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

Qy      1 MONHVLSLSAARRRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60
Db      1 MONHVLSLSAARRRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Qy      61 VEKACFMSHAIVLWKQALDEGLPYITVPEBDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120
Db      61 VEKACFMSHAIVLWKQALDEGLPYITVPEBDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

Qy      121 LETMFHNVLTSPSGVADYCGRAFPPLSEHMGTAGYIISRKAMRFFLDRAALPPGGLHP 180
Db      121 LETMFHNVLTSPSGVADYCGRAFPPLSEHMGTAGYIISRKAMRFFLDRAALPPGGLHP 180

Qy      181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALGLEHDLRLNRKQWRDS 240
Db      181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALGLEHDLRLNRKQWRDS 240

Qy      241 PANTFKHRLIRALTITKISREERKRRQREQ----FTVFPQ 275
Db      241 PANTFKHRLIRALTITKIGREERKRRQREQGLIKIIVFPQ 279

```

```

RESULT 3
US-11-102-497-6
; Sequence 6, Application US/11102497
; Publication No. US20050271690A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412

```

```

; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 6
; LENGTH: 280
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-11-102-497-6

```

```

Query Match      69.6%; Score 1007; DB 7; Length 280;
Best Local Similarity 74.7%; Pred. No. 3.8e-92;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

```

```

Qy      1 MONHVLSLSAARRRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60
Db      1 MONHVLSLSAARRRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Qy      61 VEKACFMSHAIVLWKQALDEGLPYITVPEBDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120
Db      61 VEKACFMSHAIVLWKQALDEGLPYITVPEBDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

Qy      121 LETMFHNVLTSPSGVADYCGRAFPPLSEHMGTAGYIISRKAMRFFLDRAALPPGGLHP 180
Db      121 LETMFHNVLTSPSGVADYCGRAFPPLSEHMGTAGYIISRKAMRFFLDRAALPPGGLHP 180

Qy      181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALGLEHDLRLNRKQWRDS 240
Db      181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALGLEHDLRLNRKQWRDS 240

Qy      241 PANTFKHRLIRALTITKISREERKRRQREQ 269
Db      237 RSLKVMFDLKRALGKFGREKKRKRERQ 265

```

```

RESULT 4
US-10-467-657-1326
; Sequence 1326, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1326
; LENGTH: 280
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1326

```

```

Query Match      68.9%; Score 996; DB 6; Length 280;
Best Local Similarity 74.3%; Pred. No. 4.7e-91;
Matches 200; Conservative 19; Mismatches 46; Indels 4; Gaps 1;

```

```

Qy      1 MONHVLSLSAARRRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60
Db      1 MONHVLSLSAARRRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Qy      61 VEKACFMSHAIVLWKQALDEGLPYITVPEBDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

```

```

Db      61 VERACFSHVAWLMKQALDEGLPYLAIVEDVLGKDAEKLAEADTWLEERFDKOSATVR 120
Qy      121 LETMFHVLTPSGVADYCGRAAPLLESEMGCTGYIISKRAFPFLDRPALPPEGILP 180
Db      121 LETMFHVLTPSGVADYCGRAAPLLESEMGCTGYIISKRAFPFLDRPALPPEGILP 180
Qy      181 VDLMPFSDPFRGMPYCOLNPAICAOELHYAKPHDONSALGSLIEHDLINRKQORDS 240
Db      181 VDLMPFSDPFRGMPYCOLNPAICAOELHYAKPHDONSALGSLIEHDLINRKQORDS 240
Qy      241 PANTFKRLIRALTKISREKRRQRRQ 269
Db      237 RSLKMFEDLRKALGKFERKKKMERQRRQ 265

```

RESULT 5
US-10-821-234-884
Sequence 884, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmat, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: pc_seq_genes Version 1.0
SEQ ID NO 884
LENGTH: 327
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-884

Query Match
Best Local Similarity 26.2%; Score 118.5; DB 6; Length 327;
Pred. No. 0.00052;
Matches 64; Conservative 28; Mismatches 89; Indels 63; Gaps 13;

```

Qy      5 VLSIAGAERRAHIAITFGRHGIPQFPDA---LMPSERLBQAMLEIVGSLAHNY--- 57
Db      40 VLSIARPPRRKEMTLASIMEMISGRVDAVDGMMNLSSAIRNLGVDLAPGIV-DPYSGR 98
Qy      58 -LSGVEKACFMSHVAWLMKQALDEGLPYITVFEDDVLGSGEKEFLAEDAMLOERFPDPA 116
Db      99 TLTGKVGCFLSHYSIMBEVVAAGLARVLVFEDV-----RFESEN-- 138
Qy      117 FIVLETFMTHVLTSP-SGVADYCGRAAPLLESE-----HWGTAGYTIIR 160
Db      139 FGRLEERLEMEDVAEKLISMDLITLGRKQVPEKETAIVEGLPGLVVAGYSYM-TLAVYLR 197
Qy      161 KAMRFILDRPALPPEGILPVD---LMMF-----SDPFRGMPYCOLNPAICAOE 208
Db      198 AGARKL---ASQPLRRMLPVDLFLPIMDQHPNEQYKAFWPRD-LVAFSAQPLIAA-P 252
Qy      209 LHYA 212
Db      253 THYA 256

```

RESULT 6
US-11-227-177-1
Sequence 1, Application US/11227177
Publication No. US20060014262A1
GENERAL INFORMATION:
APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
TITLE OF INVENTION: Beta 1,3-galactosyltransferase and a DNA coding for said enzyme
FILE REFERENCE: 11428WO1
CURRENT APPLICATION NUMBER: US/11/227,177

```

; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: US/10/475,587
; PRIOR FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: JP 2001-123864
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-11-227-177-1

```

Query Match
Best Local Similarity 30.2%; Score 98; DB 7; Length 481;
Pred. No. 0.092;
Matches 32; Conservative 16; Mismatches 36; Indels 22; Gaps 6;

```

Qy      29 FQFPDALMPS-ERLEQAMAEIVGSL-----AHPLYSGVEKACFMSHVAWLMKQA 76
Db      35 FQIFDALSPKHDFEKLQELVDSSNLSKDMFHSYCYQELLPQFPGCTLSHYLMKEC 94
Qy      77 LDEGLPYITVFEDDVLGSGEKEFL-ASDAMLOERFPDPAIVRL 121
Db      95 VKTNQP-VVILEDDIAL---ESNFMQALBEDCLKSPD-----FVRL 131

```

RESULT 7
US-11-087-099-2611
Sequence 2611, Application US/11087099
Publication No. US20060041961A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)B EP
CURRENT APPLICATION NUMBER: US/11/087,099
CURRENT FILING DATE: 2005-03-22
NUMBER OF SEQ ID NOS: 12464
SEQ ID NO 2611
LENGTH: 651
TYPE: PRT
ORGANISM: Brucella melitensis bioovar Abortus
US-11-087-099-2611

Query Match
Best Local Similarity 6.1%; Score 88.5; DB 7; Length 651;
Pred. No. 1.2;
Matches 50; Conservative 23; Mismatches 65; Indels 63; Gaps 11;

```

Qy      89 DVLGEGEKEFLAEDAMLOERFPDPAIVRLVLETFMTHVLTSPSGVADYCGRAF---PL 145
Db      299 DGIIGDGSSEFSGAERREIED-YQPD-----ELEKX-----RFAAASDVLGRAMEKDP 347
Qy      146 L-----ESEH---WGTAGYIISKRAFPFLDRPALP-----PGQLHFDLM 184
Db      348 IIVMGEDVIRFAGVSGF--TNALLELPDRVLAMPVLENGFTVALGAALGLRVAEI 405
Qy      185 MNSDF-----FDRGMPYCOLNPAICAOELHYAKPHDQNSALGSLIE 226
Db      406 MEGDFCFVAADQIANGISKVRMFQDGFVPVIMVEVRSFHTGYSQHSQSDSALFGMP 465
Qy      227 HDRLINRKQQRDSPPANTPKH 247
Db      466 GMRVY-----SPTNAFDY 478

```

RESULT 8
US-11-239-510-21
Sequence 21, Application US/11239510
Publication No. US20060063715A1
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
Wood, James F.
Hardman, Karl
Bird, Robert

Db 103 YYYDNDYGDYWG 114

RESULT 11

US-11-084-220-2
Sequence 2, Application US/11084220
Publication No. US20050266439A1
GENERAL INFORMATION:
APPLICANT: Hjortleifsdottir, Sigridur
APPLICANT: Thordisdottir, Audur
APPLICANT: Aevartsson, Arnthor
APPLICANT: Hreysvoldsson, Gudmundur O.
APPLICANT: Blondal, Thorarinn
TITLE OF INVENTION: Thermostable Ligase from Thermus Phage
FILE REFERENCE: 2138.001US1
CURRENT APPLICATION NUMBER: US/11/084,220
PRIOR FILING DATE: 2005-03-18
PRIOR APPLICATION NUMBER: US 60/412,461
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 395
TYPE: PRT
ORGANISM: Bacteriophage T82126
US-11-084-220-2

Query Match 5.8%; Score 84; DB 7; Length 395;
Best Local Similarity 25.6%; Pred. No. 1.7;
Matches 52; Conservative 22; Mismatches 63; Indels 66; Gaps 12;

QY 29 FQFFDLMPSEKLEQ-----MAELVPGLSAHPYLSGVKACFMGSHAV-----L 72
DB 85 FNGEELAGEAEAFKAFKRDMSWPLFVAEKVDGYLAQVYLDGGELRPSRSRSLNPPVLGAL 144
QY 73 WKQALDEGPPYIVFEDDVLLGEGEKEFLAEDA--W--LQERPDPTAFIVLETFMFMV 128
DB 145 LKRAYE-----BAMARLG---KLAAEGGRTALLLEVDPAPVWVPPYQERGVL 192
QY 129 LNSPS-GVADYC--GRAFLLES--EHW-----GTAGYIIS----- 159
DB 193 LALRSIGEGHYLLPGHFFPLPEALRYVRWEPMDPDPHFRGRIRDLQGVGCVVTDGAE 252
QY 160 ----RKAMRFLLDRFAALPREG 178
DB 253 FVFKTGMAFRLARF-LMDPEGV 274

RESULT 12

US-11-239-510-11
Sequence 11, Application US/11239510
Publication No. US20060063715A1
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
APPLICANT: Wood, James P.
APPLICANT: Hardman, Karl
APPLICANT: Bird, Robert
APPLICANT: Filpula, David
APPLICANT: Rolence, Michelle
TITLE OF INVENTION: Multivalent Antigen-Binding Proteins
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/239,510
FILING DATE: 29-Sep-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/137,297
FILING DATE: 03-May-2002
APPLICATION NUMBER: US/09/443,213
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/166,094
FILING DATE: 05-OCT-1998
APPLICATION NUMBER: US 08/392,338
FILING DATE: 22-FEB-1995
APPLICATION NUMBER: US 07/989,846
FILING DATE: 20-NOV-1992
APPLICATION NUMBER: US 07/796,936
FILING DATE: 25-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0977.003000E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-11-239-510-11

Query Match 5.8%; Score 83.5; DB 7; Length 240;
Best Local Similarity 29.8%; Pred. No. 1;
Matches 37; Conservative 19; Mismatches 37; Indels 31; Gaps 9;

QY 40 RLEQMAHEIV-GLLSAHPYLSGVKAC-----FMSHAVLW-KQALDEGLPITYV---E 88
DB 129 QIQQSDAEIVKPGAS-----VKISCRSGYTFDHALHWKQNDQGLKIVGFSPGN 181
QY 89 DVLVGEGERKELADAWLQERPDPTAFIVLETFMFMVLTSPSGVADYCGRAFLLES 148
DB 182 DLF---KTYERPKK-ATLTADKSSSTAYVQ-----LNSLTSEDSAYVFTCRS---LNM 228
QY 149 EHWG 152
DB 229 AYWG 232

RESULT 13

US-11-239-510-19
Sequence 19, Application US/11239510
Publication No. US20060063715A1
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
APPLICANT: Wood, James P.
APPLICANT: Hardman, Karl
APPLICANT: Bird, Robert
APPLICANT: Filpula, David
APPLICANT: Rolence, Michelle
TITLE OF INVENTION: Multivalent Antigen-Binding Proteins
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/239,510
FILING DATE: 29-Sep-2005
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/137,297
FILING DATE: 03-May-2002
APPLICATION NUMBER: US/09/443,213
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/166,094
FILING DATE: 05-OCT-1998
APPLICATION NUMBER: US 08/392,338
FILING DATE: 22-FEB-1995
APPLICATION NUMBER: US 07/989,846
FILING DATE: 20-NOV-1992
APPLICATION NUMBER: US 07/796,936
FILING DATE: 25-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0977.003000E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-11-239-510-19

Query Match          5.8% Score 83.5; DB 7; Length 483;
Best Local Similarity 29.8%; Pred. No.2.5;
Matches 37; Conservative 19; Mismatches 37; Indels 31; Gaps 9;

40 RLEQAMAEVL-PGLSAHPYLSGVEKAC-----FMSHAVLW-KQALDEGLPYTVF---E 88
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 130 QLOQSDALVLRPGAS-----VKISCKASGYTFDHAIHVWKQVPEQGLMWIGVSPGN 182

QY 89 DDVLIGSEEEKFLADDAVLQERFPDPTAFYRLRLEMFHVLTSRSGVADYCGRAFPILLES 148
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 183 DDF--KXNERFFKGR-ATLTADKSSSTAYVQ-----LNSLTSQSDAVYFCTRS---LNM 229

QY 149 EHWG 152
   :|:|
Db 230 AYWG 233

RESULT 14
US-10-506-454-1626
; Sequence 1626, Application US//10506454
; Publication No. US20060069386A1
GENERAL INFORMATION:
APPLICANT: Slesarev, Alexi I
APPLICANT: Mezhevaeva, Katja V
APPLICANT: Polunhin, Nikolai N
APPLICANT: Shcherbina, Olga V
APPLICANT: Shakhova, Vera V
APPLICANT: Malyykh, Andrei G
APPLICANT: Kozlyavkin, Sergei A
TITLE OF INVENTION: The Complete Genome and Protein Sequences of the Hyperthermophilic
TITLE OF INVENTION: Methanopyrus Kandleri AV19 and Monophyly of Archaeal Methanogens
FILE REFERENCE: FID001
CURRENT APPLICATION NUMBER: US/10/506,454
CURRENT FILING DATE: 2004-08-31
PRIOR APPLICATION NUMBER: PCT/US03/06664
PRIOR FILING DATE: 2003-03-04

```

```

      ; PRIOR APPLICATION NUMBER: 60/361,742
      ; PRIOR FILING DATE: 2002-03-04
      ; NUMBER OF SEQ ID NOS: 1722
      ; SOFTWARE: PatentIn version 3.2
      ; SEQ ID NO 1626
      ; LENGTH: 353
      ; TYPE: PR1
      ; ORGANISM: Methanopyrus kandleri
      ; US-10-506-454-1626

Query Match      5.7%; Score 83; DB 6; Length 353;
Beet Local Similarity 20.8%; Pred. No. 1.9; Mismatches 96; Indels 102; Gaps 13;
Matches 62; Conservative 38;

      33 DALMPSERLSQAMALVPGLSAHNYPLSGVEKACPMASHVLM-----KQALDEGLPYITVFE 88
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db DALPDHPFSEKVLNHL-----LGAABE--FVCLASPMVSPSKDLVEGVNSLRGLS 136
      89 DD-----VLLGGEKEFTLAEDWL-----QERDPP----- 114
      139 DRDLIELVLAAGNBEREVLLEMWLSLGEVNEAEAGRGREGDGIHCXYANKEKALGASWN 198
      115 --TAFIVRLFTM--PMWVLTSPSGVADYCGRAFPLLSE--HMGTAGYIISRKAMRFFLD 168
      199 LTVSIRLRRLAMREVHTIINPKADDCGYCNANTYEQVSEFNSGW-----SEAKRFFRD 251
      169 RFAALPP-----EGLHPVDLMFSD-----FFDREGMPVCOLNPAI 204
      252 EGLGPAIFEVRMWGDGDPTEVAVIYRENGERWFTVELEDNHEGFVNEREGELPVRRGVPRV 311
      205 CAQELHYAKFHDONSALGSLIENHRLNLRKQORRDSPPANPEKHLIRALTGISREKR 262
      312 HRFDLRGSRRGQ-----DRLSLRKTQAREDPGSGAGCR-----RAGREMER 353

RESULT 15
US-10-932-334-77
; Sequence 77, Application US/10932334
; Publication NO. US20050249728A1
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTI-IGF-I RECEPTOR ANTIBODY
; FILE REFERENCE: A6689
; CURRENT APPLICATION NUMBER: US/10/932,334
; CURRENT FILING DATE: 2004-09-02
; PRIOR APPLICATION NUMBER: US/10/729,441
; PRIOR FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: 10/1170,390
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 120
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic antibody structure
; US-10-932-334-77

Query Match      5.7%; Score 82; DB 6; Length 120;
Beet Local Similarity 27.6%; Pred. No. 0.59;
Matches 35; Conservative 22; Mismatches 38; Indels 32; Gaps 9

      40 RLEQAMALTV-PGSAHPYLSGVEKAC-----FMASHVLM-KQALDEGLPYITVFEEDV 91
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db QIQQSGAILMPGAS-----VKISCATGYTSSFWIEWKQRPNGGLEMI-----GEI 51
      92 LLGSG-----EKKFLAEDAMLOERPDPTAFIVRLFTMFMWVLTSPSGVADYC--GRAPPL 145
      |||:||||:||||:||||:||||:||||:||||:||||:||||:
Db LPSGCGTHYNEKFKKATFTADK--SSNTAY-----MQLSSILTSEDSAVYYCARGHSYVF 104
      146 LSEHWG 152
      :|||:|

```

Mon Apr 10 08:45:28 2006

us-09-211-691-2.rapbn

Page 7

Db 105 YDGDYWG 111

Search completed: April 7, 2006, 15:45:38
Job time : 26 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:30:46 ; Search time 573 Seconds
(without alignments)
663.242 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MONHVISLASAARRAHAD.....ISREKRRQRREQIVPQ 275

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues
Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US0606_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US0606_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US0703_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US0801_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US0801_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US0803_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US0803_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US0803_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US0806_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US0806_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US0808_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US0808_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US0809_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US0901_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US0902_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US0903_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US0904_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US0904_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US0905_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US0905_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US0907_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US0908_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US0909_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US1001_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US1001_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US1002_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US1002_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US1004_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US1005_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US1007_COMB.pep.*
38: /cgn2_6/ptodata/1/paa/US1008_COMB.pep.*
39: /cgn2_6/ptodata/1/paa/US1009_COMB.pep.*
40: /cgn2_6/ptodata/1/paa/US1010_COMB.pep.*
41: /cgn2_6/ptodata/1/paa/US1011_COMB.pep.*
42: /cgn2_6/ptodata/1/paa/US1012_COMB.pep.*
43: /cgn2_6/ptodata/1/paa/US1014_COMB.pep.*

44: /cgn2_6/ptodata/1/paa/US600_COMB.pep.*
45: /cgn2_6/ptodata/1/paa/US601_COMB.pep.*
46: /cgn2_6/ptodata/1/paa/US602_COMB.pep.*
47: /cgn2_6/ptodata/1/paa/US603_COMB.pep.*
48: /cgn2_6/ptodata/1/paa/US604_COMB.pep.*
49: /cgn2_6/ptodata/1/paa/US605_COMB.pep.*
50: /cgn2_6/ptodata/1/paa/US606_COMB.pep.*
51: /cgn2_6/ptodata/1/paa/US607_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1446	100.0	275	22	US-09-211-691-2	Sequence 2, Appl1
2	1446	100.0	275	33	US-10-317-428-2	Sequence 2, Appl1
3	1446	100.0	275	33	US-10-317-773-2	Sequence 2, Appl1
4	1288	89.1	279	36	US-10-654-528-8	Sequence 8, Appl1
5	1288	89.1	279	41	US-11-102-497-8	Sequence 8, Appl1
6	1281	88.6	279	34	US-10-467-657-1310	Sequence 1310, Ap
7	1281	88.6	279	34	US-10-467-657A-1310	Sequence 1310, Ap
8	1007	69.6	280	30	US-10-096-129-6	Sequence 6, Appl1
9	1007	69.6	280	36	US-10-654-528-6	Sequence 6, Appl1
10	1007	69.6	280	41	US-11-102-497-6	Sequence 6, Appl1
11	996	68.9	280	34	US-10-467-657-1326	Sequence 1326, Ap
12	996	68.9	280	34	US-10-467-657A-1326	Sequence 1326, Ap
13	933	64.5	312	34	US-10-472-260-176	Sequence 176, App
14	436	30.2	312	50	US-60-691-214-1389	Sequence 3189, App
15	354	25.2	255	36	US-10-603-108-2407	Sequence 2407, App
16	364	20.7	285	45	US-60-128-476-4128	Sequence 4128, App
17	300	25.2	282	50	US-60-691-214-1405	Sequence 3405, App
18	255	17.6	50	20	US-09-084-616-2	Sequence 2, Appl1
19	239	16.5	50	20	US-09-084-616-3	Sequence 3, Appl1
20	130.5	9.0	278	39	US-10-912-745-281	Sequence 281, App
21	130.5	9.0	278	39	US-10-912-745A-281	Sequence 281, App
22	130.5	9.0	278	49	US-60-512-650-418	Sequence 418, App
23	130.5	9.0	278	49	US-60-519-832-418	Sequence 418, App
24	130.5	9.0	278	49	US-60-552-390-418	Sequence 418, App
25	130.5	9.0	288	27	US-09-758-446-1206	Sequence 1206, App
26	130.5	9.0	288	32	US-10-212-759-1206	Sequence 1206, App
27	130.5	9.0	466	31	US-10-170-205E-36168	Sequence 36168, A
28	130.5	9.0	521	32	US-10-218-140-3328	Sequence 3328, App
29	130.5	9.0	622	26	US-09-611-523-272	Sequence 272, App
30	130.5	9.0	622	26	US-09-611-526-3054	Sequence 3054, App
31	130.5	9.0	622	31	US-10-170-205E-36147	Sequence 36147, A
32	130.5	9.0	622	33	US-10-305-278-272	Sequence 272, App
33	130.5	9.0	622	33	US-10-398-037-1	Sequence 1, Appl1
34	130.5	9.0	622	39	US-10-912-745-283	Sequence 283, App
35	130.5	9.0	622	39	US-10-912-745A-283	Sequence 283, App
36	130.5	9.0	622	46	US-60-236-869-1	Sequence 1, Appl1
37	130.5	9.0	622	48	US-60-453-050-9797	Sequence 9797, App
38	130.5	9.0	622	48	US-60-453-135-9797	Sequence 9797, App
39	130.5	9.0	622	48	US-60-466-412-9797	Sequence 9797, App
40	130.5	9.0	622	49	US-60-512-650-420	Sequence 420, App
41	130.5	9.0	622	49	US-60-519-832-420	Sequence 420, App
42	130.5	9.0	645	39	US-60-552-390-420	Sequence 420, App
43	130.5	9.0	645	39	US-10-912-745-280	Sequence 280, App
44	130.5	9.0	645	39	US-10-912-745A-280	Sequence 280, App
45	130.5	9.0	645	49	US-60-512-650-417	Sequence 417, App

ALIGNMENTS

RESULT 1
US-09-211-691-2
; Sequence 2, Application US/09211691
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel

APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 01957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-09-211-691-2

Query Match 100.0%; Score 1446; DB 22; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVISLASAERRRAHIDTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
DB 1 MONHVISLASAERRRAHIDTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDLLGEGEKEFLAEDAWLOERFPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDLLGEGEKEFLAEDAWLOERFPDPTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMMFSDPFDEGMPVCOINPALCAOELHYAKFHDONSALGSLIEHDLNRKQOORDS 240
DB 181 VDLMMFSDPFDEGMPVCOINPALCAOELHYAKFHDONSALGSLIEHDLNRKQOORDS 240
QY 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275

RESULT 2
US-10-317-428-2
Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000811US
CURRENT APPLICATION NUMBER: US/10/317,428
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-317-428-2

Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MONHVISLASAERRRAHIDTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60

DB 1 MONHVISLASAERRRAHIDTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
QY 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDLLGEGEKEFLAEDAWLOERFPDPTAFIVR 120
DB 61 VEKACFMSHAIVLMKQALDEGLPYITVFEDVDLLGEGEKEFLAEDAWLOERFPDPTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMMFSDPFDEGMPVCOINPALCAOELHYAKFHDONSALGSLIEHDLNRKQOORDS 240
DB 181 VDLMMFSDPFDEGMPVCOINPALCAOELHYAKFHDONSALGSLIEHDLNRKQOORDS 240
QY 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275

RESULT 3
US-10-317-773-2
Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000812US
CURRENT APPLICATION NUMBER: US/10/317,773
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-317-773-2

Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000812US
CURRENT APPLICATION NUMBER: US/10/317,773
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-317-773-2

Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000812US
CURRENT APPLICATION NUMBER: US/10/317,773
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-317-773-2

```

; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 279
; TYPE: PR1
; ORGANISM: Neisseria gonorrhoeae
US-10-654-528-8

Query Match      89.1%; Score 1288; DB 36; Length 279;
Best Local Similarity 90.0%; Pred. No. 4,2e-121;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVLSLSAARRRAHIADTFGRHGIPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVLSLSAARRRAHIATFGSRGIPPOFPDAMPSERLERAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHAIVLWKQALDBGLPYITVFEEDVLLGEGEKEFLADAWLOREPDPTAIFYR 120
DB 61 VERACFMSHAIVLWKQALDBGLPYITVFEEDVLLGEGEKEFLADAWLOREPDPTAIFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLSESHWGTAAGIISRKAMRFFLDRAVALPPEGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLSESHWGTAAGIISRKAMRFFLDRAVALPPEGLHP 180
QY 181 VDIAMFSDFFDRGMPVQCQNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDIAMFSDFFDRGMPVQCQNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRRLIRALTXTISREKRRORREQ----FIVPQ 275
DB 241 PANTFGRRLIRALTXTISREKRRORREQ----FIVPQ 275

RESULT 5
US-11-102-497-8
; Sequence 8, Application US/11102497
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
```

```

; SEQ ID NO 8
; LENGTH: 279
; TYPE: PR1
; ORGANISM: Neisseria gonorrhoeae
US-11-102-497-8

Query Match      89.1%; Score 1288; DB 41; Length 279;
Best Local Similarity 90.0%; Pred. No. 4,2e-121;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVLSLSAARRRAHIADTFGRHGIPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVLSLSAARRRAHIATFGSRGIPPOFPDAMPSERLERAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHAIVLWKQALDBGLPYITVFEEDVLLGEGEKEFLADAWLOREPDPTAIFYR 120
DB 61 VERACFMSHAIVLWKQALDBGLPYITVFEEDVLLGEGEKEFLADAWLOREPDPTAIFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLSESHWGTAAGIISRKAMRFFLDRAVALPPEGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLSESHWGTAAGIISRKAMRFFLDRAVALPPEGLHP 180
QY 181 VDIAMFSDFFDRGMPVQCQNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDIAMFSDFFDRGMPVQCQNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQQRDS 240

RESULT 6
US-10-467-657-1310
; Sequence 1310, Application US/10467657
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1310
; LENGTH: 279
; TYPE: PR1
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1310

Query Match      88.6%; Score 1281; DB 34; Length 279;
Best Local Similarity 90.0%; Pred. No. 2,2e-120;
Matches 251; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

QY 1 MONHVLSLSAARRRAHIADTFGRHGIPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVLSLSAARRRAHIATFGSRGIPPOFPDAMPSERLERAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHAIVLWKQALDBGLPYITVFEEDVLLGEGEKEFLADAWLOREPDPTAIFYR 120
DB 61 VERACFMSHAIVLWKQALDBGLPYITVFEEDVLLGEGEKEFLADAWLOREPDPTAIFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLSESHWGTAAGIISRKAMRFFLDRAVALPPEGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLSESHWGTAAGIISRKAMRFFLDRAVALPPEGLHP 180
QY 181 VDIAMFSDFFDRGMPVQCQNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDIAMFSDFFDRGMPVQCQNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQQRDS 240
```


Db 61 VEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Qy 121 LETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRFALPPEGGLHP 180
Db 121 LETMFAKVIVRPDKVLTNYENRSPFLSESHMGTAGYIISRKAMRPFLLDRFALPPEGGLHP 180
Qy 181 VDLAMFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 240
Db 181 VDLAMFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 240
Qy 241 PANTFKHRLIRALTGISREKRRORRQ 269
Db 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
RESULT 13
US-10-472-260-176 *
; Sequence 176, Application US/10472260
; GENERAL INFORMATION:
; APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY
; TITLE OF INVENTION: IMPERIAL COLLEGE INNOVATIONS LIMITED
; FILE REFERENCE: GMS/DJC/23480
; CURRENT APPLICATION NUMBER: US/10/472,260
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 176
; LENGTH: 280
; TYPE: PRN
; ORGANISM: Neisseria meningitidis (group B)
US-10-472-260-176
Query Match 64.5%; Score 933; DB 34; Length 280;
Best Local Similarity 69.5%; Pred. No. 4.1e-85;
Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;
Qy 1 MONHVISLASAARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMALVPGLSAHFYLSG 60
Db 1 MONHVISLASAARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMALVPGLSAHFYLSG 60
Qy 61 VEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Db 61 VEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Qy 121 LETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRFALPPEGGLHP 180
Db 121 LETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRFALPPEGGLHP 180
Qy 181 VDLAMFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 240
Db 181 VDLAMFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 240
Qy 241 PANTFKHRLIRALTGISREKRRORRQ 269
Db 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
RESULT 14
US-60-691-214-3189
; Sequence 3189, Application US/60691214
; GENERAL INFORMATION:
; APPLICANT: Munson, et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
; FILE REFERENCE: 28335/38815
; CURRENT APPLICATION NUMBER: US/60/691,214
; NUMBER OF SEQ ID NOS: 4387
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 3189
; LENGTH: 312
; TYPE: PRN
; ORGANISM: H. influenzae strain 86-028NP

US-60-691-214-3189
Query Match 30.2%; Score 436; DB 50; Length 312;
Best Local Similarity 35.0%; Pred. No. 1.1e-34;
Matches 100; Conservative 58; Mismatches 110; Indels 18; Gaps 5;
Qy 1 MONHVISLASAARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMALVPGLSAHFYLSG 60
Db 1 MONHVISLASAARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMALVPGLSAHFYLSG 60
Qy 51 GUSAPYISGVEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Db 51 GUSAPYISGVEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Qy 169 PDPDTAFIVRLTETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDR 169
Db 169 PDPDTAFIVRLTETMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDR 169
Qy 170 PALPPEGGLHPVDMFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 240
Db 170 PALPPEGGLHPVDMFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 240
Qy 230 ILNRKQOORDSPANTFKHRLIRALTGISREKRRORRQFIVPQ 275
Db 230 ILNRKQOORDSPANTFKHRLIRALTGISREKRRORRQFIVPQ 275
Qy 263 EKRPKIRKK---KTLKQRLTRIKENIIRALNRKKMKKQGRKEMQ 304
Db 263 EKRPKIRKK---KTLKQRLTRIKENIIRALNRKKMKKQGRKEMQ 304

RESULT 15
US-10-603-108-2407
; Sequence 2407, Application US/10603108
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: PAT93-14
; CURRENT APPLICATION NUMBER: US/10/603,108
; PRIOR FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US 09/540,263
; PRIOR FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: US 60/125,416
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2407
; LENGTH: 255
; TYPE: PRN
; ORGANISM: M. catarrhalis
US-10-603-108-2407
Query Match 25.2%; Score 364; DB 36; Length 255;
Best Local Similarity 35.2%; Pred. No. 1.7e-27;
Matches 90; Conservative 49; Mismatches 103; Indels 14; Gaps 7;
Qy 1 MONHVISLASAARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMALVPGLSAHFYLSG 60
Db 1 MONHVISLASAARRAHIAIDTFGRHGIPOFPDAMPSERLBOAMALVPGLSAHFYLSG 60
Qy 61 VEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Db 61 VEKACPMASHAVIMKQALDEGLPYITAVFEDVILGKDAEFTLADBTWLEERPDKDSAFIVR 120
Qy 120 RLETFMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRFALPPEGGLHP 179
Db 120 RLETFMFMHVLTPSPGVADYCGRAFPILSESHMGTAGYIISRKAMRPFLLDRFALPPEGGLHP 179
Qy 180 PVDLAMPFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 239
Db 180 PVDLAMPFSDPDRGMPVCOINPALCAQELHYAKFHQNSALGSLIEHDLNRKQOORDS 239
Qy 240 SPANTFKHRLIRALTGISREKRRORRQ 269
Db 240 SPANTFKHRLIRALTGISREKRRORRQ 269
Qy 231 TLAD-YGKKYRSISGK 245
Db 231 TLAD-YGKKYRSISGK 245

Mon Apr 10 08:45:28 2006

us-09-211-691-2.rapm

Page 7

Search completed: April 7, 2006, 15:41:07
Job time : 575 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:31:41 ; Search time 60 Seconds
(without alignments)
630.723 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHTSLASAEARRAHAD.....ISRRKKRQRREQIVPRQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 584368 seqs, 137612332 residues

Total number of hits satisfying chosen parameters: 584368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /SIDSS/ptocdata/1/paa/PCF_NEW COMB.pap:*
2: /SIDSS/ptocdata/1/paa/US06_NEW COMB.pap:*
3: /SIDSS/ptocdata/1/paa/US07_NEW COMB.pap:*
4: /SIDSS/ptocdata/1/paa/US08_NEW COMB.pap:*
5: /SIDSS/ptocdata/1/paa/US09_NEW COMB.pap:*
6: /SIDSS/ptocdata/1/paa/US10_NEW COMB.pap:*
7: /SIDSS/ptocdata/1/paa/US11_NEW COMB.pap:*
8: /SIDSS/ptocdata/1/paa/US60_NEW COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130.5	9.0	622	8	US-60-742-219-2224 Sequence 2224, Ap
2	113.5	7.8	527	6	US-10-513-957-29 Sequence 29, Appl
3	110	7.6	739	6	US-10-461-673-16345 Sequence 16345, A
4	90	6.2	738	6	US-11-311-940-9848 Sequence 9848, Ap
5	89	6.2	475	7	US-11-360-355-121052 Sequence 121052,
6	88.5	6.1	651	8	US-60-752-355-26732 Sequence 26732, A
7	87	6.0	771	7	US-11-360-355-133331 Sequence 133331,
8	85	5.9	1189	8	US-60-752-355-52377 Sequence 52377, A
9	84.5	5.8	386	1	PCT-US06-00964-4615 Sequence 4815, Ap
10	84.5	5.8	386	6	US-11-330-403-4815 Sequence 4815, Ap
11	84.5	5.8	951	8	US-60-752-355-16259 Sequence 16259, A
12	84	5.8	611	7	US-11-360-355-146612 Sequence 146612,
13	83.5	5.8	115	6	US-10-570-220-82 Sequence 82, Appl
14	83.5	5.8	479	8	US-60-752-355-26691 Sequence 26691, A
15	83.5	5.8	531	6	US-11-293-697-3858 Sequence 3858, Ap
16	83.5	5.8	531	6	US-60-751-420-2774 Sequence 2774, Ap
17	83.5	5.8	653	6	US-11-045-004-468 Sequence 468, App
18	83.5	5.8	803	6	US-11-045-004-2081 Sequence 2081, Ap
19	83.5	5.8	1217	6	US-10-461-673-9202 Sequence 9202, Ap
20	83.5	5.8	1336	6	US-60-742-219-1084 Sequence 1084, Ap
21	82.5	5.7	275	6	US-11-311-940-8938 Sequence 8938, Ap
22	82.5	5.7	873	8	US-60-752-355-39053 Sequence 39053, A
23	82.5	5.7	924	8	US-60-752-355-3969 Sequence 3969, Ap
24	82.5	5.7	966	8	US-60-752-355-1007 Sequence 1007, A
25	82.5	5.7	966	8	US-60-752-355-24368 Sequence 24368, A

26	82.5	5.7	966	8	US-60-752-355-38250 Sequence 38250, A
27	82	5.7	120	6	US-11-290-687-61 Sequence 61, Appl
28	82	5.7	955	8	US-60-752-355-23148 Sequence 23148, A
29	82	5.7	963	8	US-60-752-355-23441 Sequence 23441, A
30	82	5.7	1014	6	US-60-752-355-5751 Sequence 5751, Ap
31	81.5	5.6	316	6	US-10-461-673-10850 Sequence 10850, A
32	81.5	5.6	326	6	US-11-311-940-1350 Sequence 1350, Ap
33	81.5	5.6	326	8	US-60-752-355-41674 Sequence 41674, A
34	81.5	5.6	492	6	US-11-311-940-2862 Sequence 2862, Ap
35	81.5	5.6	492	8	US-60-752-355-17769 Sequence 17769, A
36	80.5	5.6	275	6	US-11-311-940-4346 Sequence 4346, Ap
37	80.5	5.6	275	6	US-11-311-940-6589 Sequence 6589, Ap
38	80.5	5.6	297	6	US-11-311-940-551 Sequence 551, App
39	80.5	5.6	473	6	US-11-311-940-5403 Sequence 5403, Ap
40	80.5	5.6	473	7	US-60-752-355-33231 Sequence 33231, A
41	80	5.5	530	7	US-11-360-355-124908 Sequence 124908,
42	80	5.5	676	8	US-60-752-355-24908 Sequence 24908, A
43	80	5.5	1583	8	US-60-752-355-17545 Sequence 17545, A
44	79.5	5.5	492	1	PCT-US06-00964-5647 Sequence 5647, Ap
45	79.5	5.5	492	6	US-11-311-940-4124 Sequence 4124, Ap

ALIGNMENTS

RESULT 1
US-60-742-219-2224
; Sequence 2224, Application US/60742219
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Cousineau, Johane
; APPLICANT: Berdewegh, Paul Van
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
; FILE REFERENCE: 306522-2000
; CURRENT APPLICATION NUMBER: US/60/742,219
; CURRENT FILING DATE: 2005-12-05
; NUMBER OF SEQ ID NOS: 7303
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2224
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-742-219-2224

Query Match 9.0%; Score 130.5; DB 8; Length 622;

Best Local Similarity 25.1%; Pred. No. 0.00096; Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;

QY 5 VISLSAAERAHADTRGRHGIPOFPDALT---MPSERLQAAELVPGISAPRY--- 57
DB 345 MINLARRQDRRRMRRLALQAOEIBCRALVEADVCKAMNTSQVVALGIQMLPGV-RDPYHGR 403
QY 58 -LSGVKACFSNHAIVLQALDEGLPIYTVPEDDV-----LLGSEBEKFLAED 104
DB 404 PLTKBELCPSFHSYVIMKREYVDRGLQKSLVFEDDDLFBIFFKRRLLMTRDVE----- 456
QY 105 AWLGRPPDPYAFIVRLTFMFWVLTSPSGVADYCGRAP---LLASEH-WGTAGYIIS 159
DB 457 ---RGGLMDLIVYRKRMQVZH---PE-----KAVPRVRLVEADYSTWTLATYIS 502
QY 160 RKAMPFLDRFAALPPEGLAPVDLMFSDPFDRECMPCQALNPAACQAEIH 210
DB 503 LQGAARKL---AABPLSKMLPVDDEL-PVMEFKH---PVSEYKAFSLNLIH 547

```
RESULT 2
US-10-513-957-29
; Sequence 29, Application US/10513957
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: KHARE, Reena
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: MAROUTIS, Joseph P.
; APPLICANT: RAMKUMAR, Jayalakshmi
; APPLICANT: CHAMPA, Narinder K.
; APPLICANT: MASON, Patricia M.
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: JIN, Pei
; APPLICANT: BECHA, Shanya D.
; APPLICANT: KABLE, Amy E.
; APPLICANT: TRAN, Oyen K.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BUREORD, Neil
; APPLICANT: GRAULI, Richard C.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: CELL ADHESION AND EXTRACELLULAR MATRIX PROTEINS
; FILE REFERENCE: PF-1415 PCT
; CURRENT APPLICATION NUMBER: US/10/513,957
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/379,840
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: US 60/381,291
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/383,183
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/394,146
; PRIOR FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7512909CD1
US-10-513-957-29

Query Match          7.8%; Score 113.5; DB 6; Length 527;
Best Local Similarity 32.6%; Pred. No. 0.03;
Matches 31; Conservative 16; Mismatches 39; Indels 9; Gaps 3;

QY 5 VISLSAARRAHIAADTFFRHGIPQFPDA---LMPSRRLQAMAEIVPGLSAHFY--- 57
DB 324 VLSLARPRRERRLASLMEISGRVVDADVDSMTLSNRNLGVLLPGIQ-DYSGR 382
QY 58 -LSGVEKACFMSHAYLWKQALDEGLPYITVFEEDV 91
DB 383 TLTKEVGCFLSHYISIMEEVVARGLARVLVFEEDV 417

RESULT 3
US-10-461-673-16345
; Sequence 16345, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Aseundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.

APPLICANT: Xue, Aidong
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Zhi Wei
APPLICANT: Yang, Yonghong
APPLICANT: Goodrich, Ryle W.
APPLICANT: Wang, Dunru
APPLICANT: Yamazaki, Victoria
APPLICANT: Ujwal, Manusha L.
APPLICANT: Ma, Yundong
APPLICANT: Chen, Rui-Hong
APPLICANT: Ghosh, Malabika
APPLICANT: Weng, Gezhi
APPLICANT: Haley-Vicente, Dana
APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: perl_genes Version 6.0
; SEQ ID NO 16345
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-673-16345

Query Match          7.6%; Score 110; DB 6; Length 739;
Best Local Similarity 24.0%; Pred. No. 0.099;
Matches 46; Conservative 34; Mismatches 82; Indels 30; Gaps 8;

QY 38 SRLROAMAEIVPG---LSAHPIYSGVEKACFMSHAYLWKQALDEGLPYITVFEEDVLL 93
DB 496 TSQLRANINEMLPGRDPPSSRLPRG--BIGGFLSHYSYVWKEVIDRELEKTLVIDDVARF 554
QY 94 GEGEEKFLAE--DANLQERFPDPTAFIVRLTFMFMHVLTSFGVADYCGRAFLLESSEH- 150
DB 555 EHQFFKKMKMKMDNIDQALDWELIYIGRKRQVKEPEKAVENVAN-----LVEADVS 607
QY 151 WGTAGAGTISRKAMRFFLRFALPPEGGLHPVDLMM-----FSDPFDREGMPVQC 199
DB 608 YWTLGIVISLGAQGLV---GANPFGKMLPVDSEPLPVMTNKHPVAEYKEIYESRDLKAFS 664
QY 200 INPALCAQELAY 211
DB 665 AEP-LLIYPTHY 675

RESULT 4
US-11-311-940-9848
; Sequence 9848, Application US/11311940
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al
```

; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)
 ; CURRENT APPLICATION NUMBER: US/11/311,940
 ; CURRENT FILING DATE: 2005-12-19
 ; NUMBER OF SEQ ID NOS: 10032
 ; SOFTWARE: Patent version 3.3
 ; SEQ ID NO 9848
 ; LENGTH: 738
 ; TYPE: PR1
 ; ORGANISM: Chlamydia trachomatis D/UW-3/CX
 US-11-311-940-9848

Query Match 6.2%; Score 90; DB 6; Length 738;

Best Local Similarity 23.4%; Pred. No. 7.4; Mismatches 64; Conservative 38; Indels 88; Gaps 18;

QY 9 ASAAERBAHIADTFGRHGIPFOFPDAMPSERL-BOAMAE-VPGISAHPIYSGVERKAC 66
 DB 139 APHACRVSIVGDGFCGWHGL-----VNPLHKVSDGQVWELFVPGILT-----GACY 183
 QY 67 MSHAVLMKQALDEGL-----PYITVF-----EDVLLGEGEKKFLAEDAMLOERF----- 111
 DB 184 K-----WENVTBSGQVLISDPYGFPGPPWSVSVVIDSYE--WTSEWLEERIKTTE 236
 QY 112 DDPTAFIVRL-----ETMFHVLTPSGVADYCGR-----APFLR-----SEHNG--T 153
 DB 237 GPMNITIEVAVGSRWEGGPELANKELADQALYCKMHTHVELLPVTEHPPLNESGYOT 296
 QY 154 AGYII-----SRKAMRFLDRFA-----LPPEGLHPVDLMFSDFFDRBGMPCV 198
 DB 297 TGYTAPTRSGVSGFEDQYFIDTWGHQHGIGVILDMWP--GHFIPDSFAMSGF--DGTPLY 351
 QY 199 Q-----LNPALCAQELHYAKFHDONSALGSLI 225
 DB 352 EYTRNPSPLPHMHTYTPYAKPEVCNPLGSLV 385

RESULT 5
US-11-360-355-121052

Sequence 121052, Application US/11360355

GENERAL INFORMATION:

APPLICANT: Boukharov, Andrey

APPLICANT: Du, Zijing

APPLICANT: Guo, Liang

APPLICANT: Kovalic, David

APPLICANT: Lu, Maolong

APPLICANT: McCarter, James

APPLICANT: Miller, Nancy

APPLICANT: Williams, Deryck

APPLICANT: Vaudin, Mark

TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS

FILE REFERENCE: 38-21(53885)

CURRENT APPLICATION NUMBER: US/11/360,355

CURRENT FILING DATE: 2006-02-24

NUMBER OF SEQ ID NOS: 171306

SEQ ID NO 121052

LENGTH: 475

TYPE: PR1

ORGANISM: Heterodera glycines

FEATURE: NAME/KEY: n1ac feature

LOCATION: (92)-(92)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 47475; Strand=-; Position=1

OTHER INFORMATION: -63,173-550,842-1061,1104-1187,1230-1552,1593-1686,1782-1819

OTHER INFORMATION: Homolog annotation: H1c-ID=CAE71465.1; Match level="QueryCoverage

OTHER INFORMATION: =89%, HitCoverage=76%, E-value=8e-61, Identity=33%; Hit descrip

OTHER INFORMATION: =Hypothetical protein CBG18383 [Caenorhabditis briggsae]

FEATURE: TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS

; OTHER INFORMATION: Essential gene: C elegans homolog=D2045.9; Phenotype=Gro, Rup, Stp
 US-11-360-355-121052

Query Match 6.2%; Score 89; DB 7; Length 475;

Best Local Similarity 23.3%; Pred. No. 5.1; Mismatches 30; Conservative 24; Indels 49; Gaps 5;

QY 2 QNHVISASAEARRHIAIDTFGRHGIPFOFPDAMPSERLBOAMAE-VPGISAHPIYSGVER 57
 DB 303 QYTLINLRRRPQKLEKREVLHLVGLDPTFRFVAVGQNTLHLSAHLRPER-DPIYK 361
 QY 58 -LSGVEKACFMSHVAVMKQALDEGLPYITVEEDVLLGEGEKKFLAEDAMLOERPDPT 115
 DB 362 RMRKGEICCFLSHFIMEIVRNHGQRAVLEDDV-----RSPNG 403
 QY 116 AIVALETM 124
 DB 404 TLILQ-KTM 411

RESULT 6

US-60-752-355-26732

Sequence 26732, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: Patent version 3.3

SEQ ID NO 26732

LENGTH: 651

TYPE: PR1

ORGANISM: Brucella melitensis biovar Abortus

US-60-752-355-26732

Query Match 6.1%; Score 88.5; DB 8; Length 651;

Best Local Similarity 24.9%; Pred. No. 8.6; Mismatches 50; Conservative 23; Indels 65; Gaps 11;

QY 89 DVLGEGEKKFLAEDAMLOERPDPTAFIVLETFMHVLTSPSGVADYCGRAF--PL 145
 DB 299 DGGILGDSGSEFGAFREIED-YQPD-----BLEKQ-----RFAAASVILGRAMKDP 347
 QY 146 L-----ESEH--WGTAGYIISRKAMRFLDRPALP-----PGLHPVDLM 184
 DB 348 IIVMGEDVHRFAGVSGF--TRNALELPDRVLAEMPIENGFTGVALGAALRGRLRVEI 405
 QY 185 MESDF-----FDRGMPVQCLNPALCAQELHYAKFHDQ--SALGSLIE 226
 DB 406 MEGDFCFVAADQIANGISKVRHMFSDGFPVIVMRVRSPTHTGSGQHSGLDPSALGMP 465
 QY 227 HDRLNRRKQQRDSPPANTFKH 247
 DB 466 GWRVY-----SPTNAVDY 478

RESULT 7

US-11-360-355-13331

Sequence 13331, Application US/11360355

GENERAL INFORMATION:

APPLICANT: Boukharov, Andrey

APPLICANT: Du, Zijing

APPLICANT: Guo, Liang

APPLICANT: Kovalic, David

APPLICANT: Lu, Maolong

APPLICANT: McCarter, James

APPLICANT: Miller, Nancy

APPLICANT: Williams, Deryck

APPLICANT: Vaudin, Mark

APPLICANT: Wu, Wei

TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS

```

; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 133331
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATRE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 59754; Strand=+; Position=1
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/10964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 4815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Rubrobacter xylophilus DSM 9941
; PCT-US06-00964-4815

Query Match
Best Local Similarity 26.1%; Score 87; DB 7; Length 771;
Matches 24; Conservative 17; Mismatches 37; Indels 14; Gaps 2;

QY 135 MCVQCLNPLCQOEHLHYAFHONSALGSLIEHDLNKKOQRDSPPANTFPHRLI---- 250
DB 579 LFTQAVEAYCADIRYANHHAEEDVQKAKHQQILERRORALDEBRO--RQKILSTR 636
QY 251 -----RALTKISREKRRORRORQFIVPF 274
DB 637 DEDROMAYREDVLALEKIROKQKQIDIVEY 668

RESULT 8
US-60-752-355-52377
; Sequence 52377, Application us/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 52377
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
; US-60-752-355-52377

Query Match
Best Local Similarity 23.7%; Score 85; DB 8; Length 1189;
Matches 75; Conservative 35; Mismatches 116; Indels 90; Gaps 19;

QY 14 REAHADTFGRHGI--PFOFPDALMPSERLQOMAEIVPGLSAHPIYSGVEXKCFMSHAYL 72
DB 176 KTAHASSP--EHGVSPMOAISLMP-----ALAEISGSPPEPDSMT---TTHAVM 223
QY 73 WKQALQ-----EGLYITVFEDVDVLAGGEKEF-----LAEDAWLQERFPDPTAFIVR 120
DB 224 GEAAFGISPADAEIYATLRTLTDDRM-----EKLCYAAISLAKKXVAGEOQLTLDIAY--- 275
QY 121 LETMFHVLUTSPSGVA-----DYCGRAFPILBSEHWGTAGYIISRKAMRFILDR 169
DB 276 -DDIFLHCENAPPAVVAHRRALDEEKI PYNSEALPWRPSEDFGRF--RAVSSSSAM--FFLGA 332
QY 170 PAALPREGLHPVDLMFSDPFD-----REGMPVCOINPALCAQEL-- 209
DB 333 GGDYP--NLNPPY---DPEPTLIGIGARIFEMRIIRNLTTRGIATTSL--PPVSATALAA 385
QY 210 --HYAKFHONSALGSLIEHDLNKKOQRDSPPAN-----TFKHRLIPALTKISR---E 259
DB 386 VARSGAHGRQPLDAVLIDPGCOAFENIGRLAI PAIKIGIATQHHQVRSIVCQRPPLVE 445
QY 260 REKRRORRORQFIVPFQ 275
```

```

DB 446 QOKR-----FRIPRR 455

RESULT 9
PCT-US06-00964-4815
; Sequence 4815, Application PC/TUS0600964
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/00964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 4815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Rubrobacter xylophilus DSM 9941
; PCT-US06-00964-4815

Query Match
Best Local Similarity 21.3%; Score 84.5; DB 1; Length 386;
Matches 47; Conservative 30; Mismatches 61; Indels 83; Gaps 10;

QY 13 ERRAHADTFGRHGIPOFPDALMPSE-----RLQOMAEIVPGLSAHPIYSGV----- 61
DB 28 ERRERLKE-----QGVRLFDGAGDREPTDPRIRALLAEAVDEVSQYPSAGSTRALREAF 83
QY 62 -----EKACFMSHAYLM-----KQALDEGLPYITVEDVYL 92
DB 84 CGMMERRRGVELDPDAEVLPAAGSKKALFHAHLPPLHHTHRRGVAYGTGPGYVERGAL 143
QY 93 LGESE---KFLAEDAWL--QERFPDPTAFIVRLETMFHVLUTSPSGVADYCGRAFPILLE 147
DB 144 PAGGALVRLVREBGFILPLEAVDPQ-----KRIIMLANTPHNTGA----- 186
QY 148 SEHWGTAGYIISRKAMRFILDRFALPREGLHPVDLMFSD 188
DB 187 -----RAPRSYLEBAAFCRE--H--DILLFSD 210

RESULT 10
US-11-330-403-4815
; Sequence 4815, Application us/11330403
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 4815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Rubrobacter xylophilus DSM 9941
; US-11-330-403-4815

Query Match
Best Local Similarity 21.3%; Score 84.5; DB 6; Length 386;
Matches 47; Conservative 30; Mismatches 61; Indels 83; Gaps 10;

QY 13 ERRAHADTFGRHGIPOFPDALMPSE-----RLQOMAEIVPGLSAHPIYSGV----- 61
DB 28 ERRERLKE-----QGVRLFDGAGDREPTDPRIRALLAEAVDEVSQYPSAGSTRALREAF 83
QY 62 -----EKACFMSHAYLM-----KQALDEGLPYITVEDVYL 92
DB 84 CGMMERRRGVELDPDAEVLPAAGSKKALFHAHLPPLHHTHRRGVAYGTGPGYVERGAL 143
QY 93 LGESE---KFLAEDAWL--QERFPDPTAFIVRLETMFHVLUTSPSGVADYCGRAFPILLE 147
DB 144 PAGGALVRLVREBGFILPLEAVDPQ-----KRIIMLANTPHNTGA----- 186
```


Qy 148 SEHMGTAGYIISRKAMFPLDRPALPPEGHLPVLDLMSFD 188
Db 187 -----RAPRSTYEEBAAPCR--H--DILDFSD 210

RESULT 11
US-60-752-16259
; Sequence 16259, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720) B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16259
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (190)..(190)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-60-752-355-16259

Query Match 5.8%; Score 84.5; DB 8; Length 951;
Best Local Similarity 22.6%; Pred. No. 34;
Matches 85; Conservative 43; Mismatches 103; Indels 145; Gaps 23;

Qy 3 NHVISLASAAE-----RRAHITAD-----TFGR--HGI---PFOFF 32
Db 62 SHMLNLANIAEIQMYRRKKAESRKGVDDEASAPTESDIDETFOQLVRLGKSKSPREVF 121
Qy 33 DALMPSERLQAMAEVLPGLSAHPYLSGVEKACFMSHAWLWKOALDEGPPYITVFEDVL 92
Db 122 DA-----LRSCQIDIV--LTAHPTQS--VRSRLQGHASL--RSLCLTQ-----L 159
Qy 93 LSEG-----BEKFLAEAMIQERPDPTAFIVRLTFMVLVSPSGVADYCGRAFPLLS 148
Db 160 CEGGSDNEEQEL--DEALQREMEMCSSGRIR-----XILAAFK--TDEIRRTPTPTPD 209
Qy 149 EHWGTAGYIISRKAMFPLDRPALPPEGHLPV-----LMMFSDPF-- 190
Db 210 E-----MRAGSYFHDITTMNGPKFLKRVDTLKNIGIDELPYAALIQFSSWNG 260
Qy 191 DREGHP-----VQQLNPALCAQELHYAKFHDONSALGSLIEHDL-----LNKK 234
Db 261 GDRGSPRYTPVETRDVCLL--ARMMAAHYFSSMADLMFELSMWRGNDELBARAMELHK 319
Qy 235 QQRR-----DSPANF-----KHRLI--RALKISREKRRQ----- 265
Db 320 SSRKAYYIASLCTPAATFWKKISPGEPYRIILGVDRLKVTYTERARQILSKGISIP 379
Qy 266 -----RRQFIVPFQ 275
Db 380 EDQTYNVQFLFPLE 395

RESULT 12
US-11-360-355-146612
; Sequence 146612, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, MaoJong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei

; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 146612
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: Coding regions on vCDNA: vCDNA=SeqID 73035; Strand=-; Position=1
; OTHER INFORMATION: -189,235-786,1093-1780,1822-2019,2063-2271
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_958786.1; Match level="QueryCovera
; OTHER INFORMATION: =87%, HitCoverage=11%, E-value=3e-15, Identity=21%; Hit descrip
; OTHER INFORMATION: =piec1 in 1 isoform 11 [Homo sapiens] gb|AA85684.1| piec11 11 [

US-11-360-355-146612

Query Match 5.8%; Score 84; DB 7; Length 611;
Best Local Similarity 25.0%; Pred. No. 21;
Matches 23; Conservative 18; Mismatches 37; Indels 14; Gaps 2;

Qy 195 MPVQQLNPALCAQELHYAKFHDONSALGSLIEHDLNKKQORRDSPPANTFKHRLI----- 250
Db 419 LPTROAVEAYCADIRYANHHAEDEVQKATKEHQQILEKRRALBBERO--RQKLSTWR 476
Qy 251 -----RALKISREKRRQRRRQFIVPF 274
Db 477 DEDRWAYREDPVLALEKIRQKQKQIDVEY 508

RESULT 13
US-10-570-220-82
; Sequence 82, Application US/10570220
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Kashmiri, Syed V.S.
; APPLICANT: Schlom, Jeffrey
; APPLICANT: Padlan, Eduardo A.
; TITLE OF INVENTION: MINIMALLY IMMUNOGENIC VARIANTS OF SDR-GRAFTED HUMANIZED ANTIBODY
; FILE REFERENCE: CC49 AND THEIR USE
; CURRENT APPLICATION NUMBER: US/10/570,220
; CURRENT FILING DATE: 2006-02-28
; PRIOR APPLICATION NUMBER: 60/498,903
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: PCT/US04/028004
; PRIOR FILING DATE: 2004-08-27
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-570-220-82

Query Match 5.8%; Score 83.5; DB 6; Length 115;
Best Local Similarity 29.8%; Pred. No. 2.6;
Matches 37; Conservative 19; Mismatches 37; Indels 31; Gaps 9;

Qy 40 RLQAMAEIV--PGLSAHPYLSGVEKAC-----FMSHVLW--KOALDEGLPYITVF--E 88
Db 3 QUGQSDAHLYKRGAS-----VKISCKASGTFPTDHAIIHWKQNPQGLIEWIGTFSPGN 55
Qy 89 DDVLIGSGEKEFLAEDAWLQERFPDPTAFIVRLTFMVLVSPSGVADYCGRAFPLLS 148
Db 56 DDF--KYMERFPKGR--ATLTADKSSSTAYVQ-----INSILTSBDSAVYFCTRS---LNM 102
Qy 149 EHWG 152

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:52 ; Search time 280.271 Seconds
(without alignments)
8315.447 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 1 cttgaggagctcatatgtaaa.....acaaatattgcggtttac 41

Sequence: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Scoring table: 5883141 seqs, 28421725653 residues

Searched: Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBdb1.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_scs.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	70.7	1875	1 U04328	U04328 Neisseria m
2	29	70.7	4176	1 NM85YN	X78068 N.meningit
3	29	70.7	4908	1 NM95053	M95053 Neisseria m
4	29	70.7	110000	1 AE002098	AE002098 Neisseria
5	29	70.7	349980	6 AX043922	AX043922 Sequence
6	28	68.3	687	1 NM060146	U60146 Neisseria m
7	25.8	62.9	224302	14 AC103481	AC103481 Rattus no
8	25.4	62.0	41	6 AR271714	AR271714 Sequence
9	25.4	62.0	41	6 AR481803	AR481803 Sequence
10	25.4	62.0	41	6 AR527402	AR527402 Sequence
11	25.4	62.0	41	6 AR609681	AR609681 Sequence
12	25.4	62.0	41	6 AX934469	AX934469 Sequence
13	25.2	61.5	10029	1 AE015576	AE015576 Shewanell
14	25.2	61.5	301439	1 AE015943	AE015943 Clostridi
15	24.8	60.5	139325	8 AL589943	AL589943 Human DNA
16	24.8	60.5	165199	8 AC003973	AC003973 Homo sapi
17	24.8	60.5	252718	9 AC116106	AC116106 Mus muscu
18	24.6	60.0	59915	14 AC100984	AC100984 Mus muscu

C	19	24.6	60.0	110000	14 BX890561_1	Continuation (2 of
C	20	24.6	60.0	131209	8 AC126177	AC126177 Homo sapi
C	21	24.6	60.0	141017	9 AC126444	AC126444 Mus muscu
C	22	24.6	60.0	180932	9 AC131989	AC131989 Mus muscu
C	23	24.6	60.0	191606	9 AC126272	AC126272 Mus muscu
C	24	24.4	59.5	2966	6 CQ714591	CQ714591 Sequence
C	25	24.4	59.5	4207	8 AF061935	AF061935 Homo sapi
C	26	24.4	59.5	4356	8 BC022792	BC022792 Homo sapi
C	27	24.4	59.5	5984	8 AB018343	AB018343 Homo sapi
C	28	24.4	59.5	168499	8 AC092037	AC092037 Homo sapi
C	29	24.2	59.0	720	10 BV655397	BV655397 S216P6047
C	30	24	58.5	110000	1 CP000020_13	Continuation (14 o
C	31	24	58.5	174375	9 AC124422	AC124422 Mus muscu
C	32	24	58.5	241329	14 AC106206	AC106206 Rattus no
C	33	23.8	58.0	152095	9 AC154271	AC154271 Mus muscu
C	34	23.8	58.0	167612	14 AC119140	AC119140 Rattus no
C	35	23.8	58.0	180558	9 AC121840	AC121840 Mus muscu
C	36	23.8	58.0	204672	9 AC122831	AC122831 Mus muscu
C	37	23.8	58.0	212749	14 AC115435	AC115435 Rattus no
C	38	23.6	57.6	50083	8 AL359708	AL359708 Human DNA
C	39	23.6	57.6	110000	1 CR522870_25	Continuation (26 o
C	40	23.6	57.6	148965	14 AC136069	AC136069 Rattus no
C	41	23.6	57.6	180745	9 AC123033	AC123033 Mus muscu
C	42	23.6	57.6	196791	14 AC118298	AC118298 Rattus no
C	43	23.4	57.1	4842	9 BC065119	BC065119 Mus muscu
C	44	23.4	57.1	4993	9 AK122372	AK122372 Mus muscu
C	45	23.4	57.1	12037	1 AE013886	AE013886 Versinia

ALIGNMENTS

RESULT 1	U04328	1875 bp DNA linear BCT 15-APR-1994
LOCUS	U04328	Neisseria meningitidis NMB CMP-N-acetylneuraminic acid synthetase
DEFINITION	U04328	(synb) and synx (synx) genes, complete cds.
ACCESSION	U04328	
VERSION	U04328.1	GI:460144
KEYWORDS		
SOURCE		Neisseria meningitidis
ORGANISM		Neisseria meningitidis
REFERENCE		Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
AUTHORS		1 (bases 1 to 1875)
TITLE		Swartley, D.S. and Stephens, D.S.
JOURNAL		Identification of a genetic locus involved in the biosynthesis of N-acetyl-D-mannosamine, a precursor of the (alpha 2->8)-linked polysialic acid capsule of serogroup B Neisseria meningitidis
PUBMED		J. Bacteriol. 176 (5), 1530-1534 (1994)
REFERENCE		8113198
AUTHORS		2 (bases 1 to 1875)
TITLE		Stephens, D.S.
JOURNAL		Direct Submision
FEATURES		Submitted (13-DEC-1993) David S. Stephens, Emory University, Department of Medicine, 69 Butler Street, Atlanta, GA 30303, USA
source		Location/Qualifiers
		1..1875
		/organism="Neisseria meningitidis"
		/mol_type="unassigned DNA"
		/strain="NMB"
		/isolate="N7"
		/db_xref="taxon:487"
		/note="serogroup B"
		52..1185
		/gene="synx"
		52..1185
		/gene="synx"
		/function="proposed to be involved in the biosynthesis of N-acetyl-D-mannosamine"
		/codon_start=1
		/evidence=experimental
		/transl_table=11

```
/product="synx"
/protein_id="AA17654.1"
/db_xref="GI:460145"
/translation="MKRLICTGTTRADFGKLPKLLAYIENHPDELHLIVTGMHMKRT
YGRTYKVTRENYQHTYLFNSQIQEPPGAVLGNITPISRLSDEIEPDWMHIGDRL
EALGAAGVAGLSRLVCHIEGELSGTVDSIRHSISKLSHILVANEQAVTRLVQMG
EKRIHIIIGSPDLVMASTLPSELEKVEYGLPYENYGISMFHPTTEAHLMPOYA
AQPFALELSCGNIISYIPNDGTESILOELKQSDKFIAPFPIRPEYFVLKKA
KFWGNSAGIREAPLYGVPSIDVGTQSNRHMGSIIHTDYETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETNVSAQKRFIDNL"
1189..1875
/gene="synb"
1189..1875
/gene="synb"
*/standard_name="CMP-NANA synthetase"
/codon_start=1
/translation="1"
/product="CMP-N-acetylneuraminic acid synthetase"
/protein_id="AA17655.1"
/db_xref="GI:460146"
/translation="MEKONIAVILARQNSKGLPKNLKRNNGISILGHTINNAISKC
PRLIVSTGGLIABEAKNGVEVLRPAELASPTASSISGVIALETIGNSGVTL
LQPTSPRTGAHIREAPFLPDEKIRGSVSAQPMHHPKTLLOINNGEYAPMRHLS
LEQPRQQLPQAFRFGALYINDTASLIANNCFIAPTKLYIMSHQSDSIDIDTELQDQ
AENILNHKS"
```

ORIGIN

```
Query Match      70.7%; Score 29; DB 1; Length 1875;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      13 TATGGAACAAATATTCGGTTATAC 41
      |||||
Db      1188 TATGGAACAAATATTCGGTTATAC 1216
```

```
RESULT 2
NMBSYN      4176 bp      DNA      linear      BCT 18-APR-2005
LOCUS      N.meningitidis (group B) ctra, synb, synb, sync and
DEFINITION      sialyltransferase genes.
ACCESSION      X78068
VERSION      X78068.1 GI:530037
KEYWORDS      CMP-NeuNAC synthetase; ctra gene; sialyltransferase; synb gene;
SOURCE      Neisseria meningitidis serogroup B
ORGANISM      Neisseria meningitidis serogroup B
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE      1
AUTHORS      Ganguli,S., Zapata,G., Wallis,T., Reid,C., Boulnois,G., Vann,W.F.
and Roberts,I.S.
TITLE      Molecular cloning and analysis of genes for sialic acid synthetase
in Neisseria meningitidis group B and purification of the
meningococcal CMP-NeuNAC synthetase enzyme
JOURNAL      J. Bacteriol. 176 (15), 4583-4589 (1994)
PUBMED      8045888
REFERENCE      2 (bases 1 to 4176)
AUTHORS      Ganguli,S.
TITLE      Direct Submission
JOURNAL      Submitted (09-MAR-1994) S. Ganguli, University of Leicester, Dept
of Microbiology, PO Box 138, Medical Sciences Building, University
Road, Leicester LE1 9HN, UK M64389.
COMMENT      Related sequences: M57677 & M64389.
FEATURES
source      1..4176
location/Qualifiers
1..4176
/organism="Neisseria meningitidis serogroup B"
/mol_type="genomic DNA"
/isolate="NCTC8249"
/db_xref="taxon:491"
/complement(<1..657)
/gene="ctrA"
/mRNA      complement(<1..657)
```

```
gene
792..1925
/gene="synb"
792..1925
/gene="synb"
/codon_start=1
/translation="1"
/protein_id="CAA54982.1"
/db_xref="GI:530038"
/db_xref="GOA:O57141"
/db_xref="InterPro:IPR003331"
/db_xref="UniProt/TREMBL:O57141"
/translation="MKRLICTGTTRADFGKLPKLLAYIENHPDELHLIVTGMHMKRT
YGRTYKVTRENYQHTYLFNSQIQEPPGAVLGNITPISRLSDEIEPDWMHIGDRL
EALGAAGVAGLSRLVCHIEGELSGTVDSIRHSISKLSHILVANEQAVTRLVQMG
EKRIHIIIGSPDLVMASTLPSELEKVEYGLPYENYGISMFHPTTEAHLMPOYA
AQPFALELSCGNIISYIPNDGTESILOELKQSDKFIAPFPIRPEYFVLKKA
KFWGNSAGIREAPLYGVPSIDVGTQSNRHMGSIIHTDYETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETNVSAQKRFIDNL"
1929..2615
/gene="synb"
1929..2615
/gene="synb"
/codon_start=1
/translation="1"
/product="CMP-NeuNAC synthetase"
/protein_id="CAA54983.1"
/db_xref="GI:530039"
/db_xref="GOA:POA027"
/db_xref="InterPro:IPR003329"
/db_xref="UniProt/Swiss-Prot:POA027"
/translation="MEKONIAVILARQNSKGLPKNLKRNNGISILGHTINNAISKC
PRLIVSTGGLIABEAKNGVEVLRPAELASPTASSISGVIALETIGNSGVTL
LQPTSPRTGAHIREAPFLPDEKIRGSVSAQPMHHPKTLLOINNGEYAPMRHLS
LEQPRQQLPQAFRFGALYINDTASLIANNCFIAPTKLYIMSHQSDSIDIDTELQDQ
AENILNHKS"
```

```
gene
CDS
```

```
gene
CDS
```

```
CDS
```

```
stem_loop
ORIGIN
```

```
Query Match
```

```
/gene="ctrA"
792..1925
/gene="synb"
792..1925
/gene="synb"
/codon_start=1
/translation="1"
/protein_id="CAA54982.1"
/db_xref="GI:530038"
/db_xref="GOA:O57141"
/db_xref="InterPro:IPR003331"
/db_xref="UniProt/TREMBL:O57141"
/translation="MKRLICTGTTRADFGKLPKLLAYIENHPDELHLIVTGMHMKRT
YGRTYKVTRENYQHTYLFNSQIQEPPGAVLGNITPISRLSDEIEPDWMHIGDRL
EALGAAGVAGLSRLVCHIEGELSGTVDSIRHSISKLSHILVANEQAVTRLVQMG
EKRIHIIIGSPDLVMASTLPSELEKVEYGLPYENYGISMFHPTTEAHLMPOYA
AQPFALELSCGNIISYIPNDGTESILOELKQSDKFIAPFPIRPEYFVLKKA
KFWGNSAGIREAPLYGVPSIDVGTQSNRHMGSIIHTDYETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETNVSAQKRFIDNL"
1929..2615
/gene="synb"
1929..2615
/gene="synb"
/codon_start=1
/translation="1"
/product="CMP-NeuNAC synthetase"
/protein_id="CAA54983.1"
/db_xref="GI:530039"
/db_xref="GOA:POA027"
/db_xref="InterPro:IPR003329"
/db_xref="UniProt/Swiss-Prot:POA027"
/translation="MEKONIAVILARQNSKGLPKNLKRNNGISILGHTINNAISKC
PRLIVSTGGLIABEAKNGVEVLRPAELASPTASSISGVIALETIGNSGVTL
LQPTSPRTGAHIREAPFLPDEKIRGSVSAQPMHHPKTLLOINNGEYAPMRHLS
LEQPRQQLPQAFRFGALYINDTASLIANNCFIAPTKLYIMSHQSDSIDIDTELQDQ
AENILNHKS"
2616..3665
/gene="sync"
2616..3665
/gene="sync"
/codon_start=1
/translation="1"
/protein_id="CAA54984.1"
/db_xref="GI:530040"
/db_xref="GOA:Q7DDU0"
/db_xref="InterPro:IPR004144"
/db_xref="InterPro:IPR006014"
/db_xref="InterPro:IPR006190"
/db_xref="UniProt/TREMBL:Q7DDU0"
/translation="WONNNEPKIGNSGVNHEPILICEIGINHRGSLKTAPEWDA
INAGAVNQHIVYEDMSDAKQVIFGNADVSYIEMRCALNEDEIYLKXEVES
KGMIFISTPSPKAAALRLQKMDIPAYKISGECNNYPLIKLVASFGPFIILSTGNST
BSIKSVISIRAGVVALHCTNLYPPYEDVLRGANNDSSEAFPDATIGLSDHTLD
NYACGAAVAGLSRLVCHIEGELSGTVDSIRHSISKLSHILVANEQAVTRLVQMG
EKRIHIIIGSPDLVMASTLPSELEKVEYGLPYENYGISMFHPTTEAHLMPOYA
AQPFALELSCGNIISYIPNDGTESILOELKQSDKFIAPFPIRPEYFVLKKA
KFWGNSAGIREAPLYGVPSIDVGTQSNRHMGSIIHTDYETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETNVSAQKRFIDNL"
3749..54176
/codon_start=1
/translation="1"
/product="sialyltransferase"
/protein_id="CAA54985.1"
/db_xref="GI:530041"
/db_xref="GOA:O79AE7"
/db_xref="InterPro:IPR010866"
/db_xref="UniProt/TREMBL:O79AE7"
/translation="MKRIKALFQKRFQDSMULTPSPYLTTPRNNLPIVSNLQ
LNQVSLIKQLTNLVLITYSKLNKMPKVHQSANKLFEESITLDELPSPNNT
PKLLIYRSYKKIINIIPALYMLSFQHSYSLISAK"
3773..3790
/notes="putative"
70.7%; Score 29; DB 1; Length 4176;
```

Best Local Similarity 100.0%; Pred. No. 20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAACAATAATTCGGTTATAC 41
Db 1928 TATGAAAAACAATAATTCGGTTATAC 1956

RESULT 3
LOCUS NM95053

DEFINITION NM95053 4908 bp DNA linear BCT 04-Apr-2002
Neisseria meningitidis siala (siala), CMP-NeuNac synthetase (sialb),
sialc (sialc), and siald (siald) genes, complete cds.

ACCESSION NM95053 M64289 M96562
VERSION M95053.1 GI:520732
SOURCE siala, sialb, sialc, siald.
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

REFERENCE

AUTHORS

TITLE

COMMENT

FEATURES

source

-10_signal

gene

CDS

1. .4908
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/isolate="B1940"
/specific_host="Homo sapiens"
/db_xref="taxon:487"
51. .56
/standard_name="-10 signal"
174. .1307
/gene="siala"
174. .1307
/gene="siala"
/codon_start=1
/transl_table=11
/product="Siala"
/protein_id="AAA20475.1"
/db_xref="GI:520732"
/translation="MKRIICITGTRADFGKLPKLAIIENHPDLIELIVTGMMNKT
YKATREVTRENYQHTYLFENIQGSPMGAVLNTITPISRDELEPDMVMTHGRL
EALAGAVALISRLVCHIEGELSGVDSIRHSISKSHILVANEOAVTLLVQMG
EKKHIIIGSPDLVDWASSTPLSLEVEKYGLPYENYGISMFHVTTEAHLMPPQA

gene

CDS

gene

CDS

gene

CDS

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 19;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAACAATAATTCGGTTATAC 41
Db 1310 TATGAAAAACAATAATTCGGTTATAC 1338

RESULT 4

AE002098 00/c

WPCOMMENT

Sequence split into 23 fragments LOCUS AE002098 Accession AE002098
Fragment Name Begin End
AE002098_00 1 110000
AE002098_01 100001 210000
AE002098_02 200001 310000
AE002098_03 300001 410000

AOYKALTELSONIISIPYNDGTESILOELLKYODKFIAPFSIRFEYVLTKHA
KPMVNSAGIREAPLYGVPSIDVGTRONNRMHGXIIHDEYTKNIIPALQACSLG
KREADPTNGGDRISTERTERFAEYINNPEYNNWSAQRFDIANN"
1311. .1997
/gene="sialb"
1311. .1997
/gene="sialb"
/citation=[2]
/codon_start=1
/transl_table=11
/product="CMP-NeuNac synthetase"
/protein_id="AAA20476.1"
/db_xref="GI:520734"
/translation="MEKONIAVILARONSGULPIKULRRMNGISLGHITNAISSKC
PRLIYSTPGGLIABEAKRFGYVULRPRALASDTPASSISGVYHALETGNSGTYTL
LQPTSPRTGAHIREAFSLPDEKIGSVVACWEMHPLKTLQINNGAPRHLSD
LEQPRQOLQAFRPNGAIIYINDYASLIANNCFIAPTKYIMSHQSDIDIDELDQQ
ANILNHKES"
1998. .3047
/gene="sialc"
1998. .3047
/gene="sialc"
/citation=[1]
/codon_start=1
/transl_table=11
/product="sialc"
/protein_id="AAA20477.1"
/db_xref="GI:520735"
/translation="MQNNNEFKIGNRVGVNHPPLICIGINHGSLKTAPEMVDAA
VNAGAEVGHQTHIVDEMSDEAKQVTPNAAVSIYEIMERCALNEDBIKLEAYES
KCMIFSTPSPSAAALRIORDIIPAKISGCONNPLIKIVASPEKPIIISGMMISI
BSIKSVETIRERAGVPYVALLHCTNIYPTPEYRRCGMNDLSBAPDAIIGSDHLD
NYACGAVAGSISIERHFTDMRDPDPIVCSMPDITKELQGHAKLALAGGKGD
TIIAGKPTKDPAPASVAVDKOI KGGELLSGNLWKRPNQDPSVNEYETLPGKVA
CNIRKAQOIKTKDIE"
3131. .4618
/gene="siald"
3131. .4618
/gene="siald"
/citation=[3]
/codon_start=1
/transl_table=11
/protein_id="AAA20478.1"
/db_xref="GI:520736"
/translation="MLKTIKALPQPKKFFQDSMWLTSPFYLTPPRNLFVINSIQ
LNOVSLIKIQLTNLIVLYTSKMLKMPKIVHOSANNLFESIYLFELPRSPNIT
PKKLIYIYSYKKIINIOPALYMSFTGHYSIJSIAKKNIITHLIDEGTYAP
LLESFSPYTKLERYLIGNLNIKGYIDHFDIILVPPRYAKKIIPAKKYNRPFAHG
GISINNINIANLKQYQISKNQYIPVSORPISDDLKYKSIIVILNISLSOIGKFIK
LHPKEMGNVYMSLPLNWEINPRIVYINPEPFLIPLIYLTNPKQIISGLASSLIYT
PLSPSTOCLSTIGELIINLIQKTSVNEENTEMIQEHLEITIKKRNFINLNDLNGVISNP
LFRTEBTEFTLLKSAEPAYKSNYFOALFYQGLASKNNTIAGHAKLWYNNALYNNQO
IYMEYSIDIFYIDNISVDPSHSDKDLTKWIKIKHYYSADNRIGDR"

LOCUS	AE002098_04	400001	510000		
DEFINITION	Neisseria meningitidis MC58	500001	610000		
ACCESSION	AE002098_05	600001	710000		
VERSION	AE002098_06	700001	810000		
KEYWORDS	AE002098_07	800001	910000		
SOURCE	AE002098_08	900001	1010000		
ORGANISM	AE002098_09	100001	1110000		
REFERENCE	AE002098_10	1100001	1210000		
AUTHORS	AE002098_11	1200001	1310000		
	AE002098_12	1300001	1410000		
	AE002098_13	1400001	1510000		
	AE002098_14	1500001	1610000		
	AE002098_15	1600001	1710000		
	AE002098_16	1700001	1810000		
	AE002098_17	1800001	1910000		
	AE002098_18	1900001	2010000		
	AE002098_19	2000001	2110000		
	AE002098_20	2100001	2210000		
	AE002098_21	2200001	2272360		
	AE002098_22	2272360	2272360		
LOCUS	AE002098	2272360	bp	DNA	circular BCT 26-MAY-2005
DEFINITION	Neisseria meningitidis MC58				complete genome.
ACCESSION	AE002098				AE002359-AE002564
VERSION	AE002098.2				GI:66731897
KEYWORDS					
SOURCE	Neisseria meningitidis MC58				
ORGANISM	Neisseria meningitidis MC58				
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.				
AUTHORS	1 (bases 1 to 2272360)				
	Nelsson, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Dodson, R.U., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Parksey, D.S., Blair, E., Ciltone, H., Ciltone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V., Pizsa, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappoli, R. and Venter, J.C.				
	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58				
TITLE	Science 287 (5459), 1809-1815 (2000)				
JOURNAL	10710307				
PUBMED	2 (bases 1 to 2272360)				
REFERENCE	Tetzelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Ciecko, A., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T.M., Parksey, D.S., Blair, E., Ciltone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H.M., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V., Deboy, R.T., Pizsa, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappoli, R. and Venter, J.C.				
AUTHORS	Direct Submission				
TITLE	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712				
JOURNAL	Medical Center Dr. Rockville, MD 20850, USA				
REFERENCE	3 (bases 1 to 2272360)				
AUTHORS	Tetzelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Ciecko, A., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T.M., Parksey, D.S., Blair, E., Ciltone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H.M., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V., Deboy, R.T., Pizsa, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappoli, R. and Venter, J.C.				
TITLE	Submitted (18-MAY-2005) The Institute for Genomic Research, 9712				
JOURNAL	Medical Center Dr. Rockville, MD 20850, USA				
REFERENCE	Sequence update by submitter				
REMARK	On or before May 26, 2005 this sequence version replaced				
COMMENT	gi:7225269, gi:7225241, gi:7225245, gi:7225252, gi:7225266, gi:7225629, gi:7225277, gi:7413418, gi:7413419, gi:7225307, gi:7225314, gi:7225327, gi:7225337, gi:7413420, gi:7413421,				

FEATURES	source	gene	CDS	gene	CDS
1..2272360	Location/Qualifiers				
1..2272360	location="Neisseria meningitidis MC58"				
1..2272360	mol_type="genomic DNA"				
1..2272360	strain="MC58"				
1..2272360	db_xref="taxon:122586"				
1..2272360	note="serogroup: B"				
1..2272360	complement(7..498)				
1..2272360	locus_tag="NM00001"				
1..2272360	complement(7..498)				
1..2272360	locus_tag="NM00001"				
1..2272360	note="identified by similarity to EGAD:160712; match to protein family HMM PF00583"				
1..2272360	codon_start=1				
1..2272360	transl_table=1				
1..2272360	product="putative acetyltransferase"				
1..2272360	protein_id="AAF0480.1"				
1..2272360	db_xref="GI:722526"				
1..2272360	translation="MNSLFDVNTVYFTRKAGHIGRLVOALFEEMHGFEPMSSVDKIHAYGGLCLMDDELPAFAVVDGSLIGLSAAVQAFHDMSEFPYEVWLGVFVLPYRGGGIGRLVHACIGARSLGIKFLYTPDVOQFYFSEFVAVVGRHRHNGEMVTVMRLDV DKN"				
1..2272360	complement(502..897)				
1..2272360	locus_tag="NM00002"				
1..2272360	complement(502..897)				
1..2272360	locus_tag="NM00002"				
1..2272360	note="identified by Glimmer2; putative"				
1..2272360	codon_start=1				
1..2272360	transl_table=1				
1..2272360	product="hypothetical protein"				
1..2272360	protein_id="AAF0481.1"				
1..2272360	db_xref="GI:722527"				
1..2272360	translation="MPSDVGIRLQTPAFKMLKMKKIFFLWVVFSTSWAGDAEDNLLSIGSVYRLALQKNNLIDGKITGMQSDDEADARRRLQTAQADARAEAEIPAMAAQKARQ				

gene AEDLRQIGVRLDHANNAVYGAGCTKASGN"
complement(918..2312)
/gene="gltx"
/locus_tag="NMB0003"
complement(918..2312)
/gene="gltx"
/locus_tag="NMB0003"
/EC_number="6.1.1.17"
/note="identified by similarity to BGAD:16197; match to protein family HMM PF00743; match to protein family HMM TIGR00464"
/codon_start=1
/transl_table=11
/product="glutamyl-tRNA synthetase"
/protein_id="AAF04082.1"
/db_xref="GI:7225228"
/translation="MTVKTFRPAPSPPTGYTHIGVTRTALPSMAFARHKKPEFLRIEDTDLARSTAESVNIIIDGMKRVGLNVDNADNVVYQTRRFRDKEVIAELERKGAHYCYC
SKEELAMREKAKESGATYDRRMRPEVGTLEIPSDVQPVVRFTPTDGVTKMTDL
VKGESIIIPNALDLDLIARADGPTVNFQVVDYDMGVTHVIRGDDHNNPDKOINI
LKAIDANLPEYGHLPWILNBOGKKSISRESDPTAITDQGMGILPREAMNTYLARLGMA
HGDDEFTMEQPIEMFDLKDVSPPSRMDLKLWYNGEHKITTPGKLAELVPRILA
LNDIHTEKPALEDVLELVKDRPDNTLADCEFYFYVQTPAEADVQKMDSEAR
MLRFAERLEGLDPMNAEAIHDLFPKPCDEBEGIMGKLPRLAVCGTAKTSVDVAVL
ALIGKEVVKRIRA"
complement(2517..3161)
/locus_tag="NMB0004"
complement(2517..3161)
/locus_tag="NMB0004"
/note="identified by similarity to EGAD:132655; match to protein family HMM PF02698"
/codon_start=1
/transl_table=11
/product="BpiH/SdmH-related protein"
/protein_id="AAF04083.1"
/db_xref="GI:7225229"
/translation="NNKRLFCSRNKGRLRYLLGGFCLSVPELLVPASSVMAVYRTGQ
VLPVYRADAAVLGAAMDKRSPVFERIRINAIALYOSRVRGKIVFGSGTKGYM
TAAVGRYRALKOGVPARNILPENTSRNTYENINRPLRLANGASVAYIVSDPHLA
RAABMAEDGVRYNMSATPTTTPDAKNKKIKFLOGVALSLTRLEKMGDRDFPN"
complement(3158..3511)
/gene="arbc"
/locus_tag="NMB0005"
complement(3158..3511)
/gene="arbc"
/locus_tag="NMB0005"
/note="identified by similarity to EGAD:124632; match to

Query Match 70.7%; Score 29; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGGAAAAACAATAATTCGGTTATAC 41
|||||
Db 79111 TATGGAAAAACAATAATTCGGTTATAC 79083

RESULT 5
AX043922 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 1 from Patent WO0066791.
DEFINITION AX043922
ACCESSION AX043922
VERSION AX043922.1 GI:11342850
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.

REFERENCE
1 Piza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Masiagnani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scarlato,V., Rappuoli,R., Frazzer,C.M. and Grandi,G.
Neisseria genomic sequences and methods of their use

TITLE

JOURNAL Patent: WO 0066791-A 1 09-NOV-2000;
CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Source
Location/Qualifiers
1..349980
/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.
seq 1: 1 to 349980 349980 bases
seq 108: 300001 to 649980 349980 bases
seq 109: 600001 to 949980 349980 bases
seq 110: 900001 to 1249980 349980 bases
seq 111: 1200001 to 1549980 349980 bases
seq 112: 1500001 to 1849980 349980 bases
seq 113: 1800001 to 2149980 349980 bases
seq 114: 2100001 to 2272325 172325 bases"

ORIGIN

Query Match 70.7%; Score 29; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGGAAAAACAATAATTCGGTTATAC 41
|||||
Db 79104 TATGGAAAAACAATAATTCGGTTATAC 79076

RESULT 6
NM060146 687 bp DNA linear BCT 11-JUN-1997
LOCUS Sequence 1 from Patent WO0066791.
DEFINITION NM060146
ACCESSION NM060146
VERSION NM060146.1 GI:1549340
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 687)
Gibbert,M., Watson,D.C. and Wakarchuk,W.W.
Purification and characterization of the recombinant CMP-sialic
acid synthetase from Neisseria meningitidis
Biotechnol. Lett. 19, 417-420 (1997)
2 (bases 1 to 687)
Gibbert,M., Watson,D.C. and Wakarchuk,W.W.
Direct Submision
Submitted (05-JUN-1996) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada

FEATURES
source
Location/Qualifiers
1..687
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="405Y; NRC4030"
/db_xref="taxon:487"
/note="capsule type: Y; LPS type: L3"
1..687
/note="first 40 residues of the N-terminus of the
recombinant product were determined experimentally."
/codon_start=1
/evidence="experimental"
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AAB60780.1"
/db_xref="GI:1549341"
/translation="MEKONIAVILARONSKGLPIKNLRKNGGISLGHITVAATSSK
FDRITVDSGLIABRANRQVAVVLRASLADTSISGYTHALETTGNSGTVTL
LQPTSLKTHIRAFSLFDKIKGVSAACPMHHPLKTLTLQINNEVAPMRSLD
LEPQKQLPQAFRPGAIYINDTASLIANNCFIAPTLTYIMSHODSIDITDELQO
AENILNHRKS"

ORIGIN

Query Match 68.3%; Score 28; DB 1; Length 687;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

14 ATGGAAAAACAATATTTGGGTATAC 41
 1 ATGGAAAAACAATATTTGGGTATAC 28

Db

RESULT 7
 AC103481
 LOCUS Rattus norvegicus clone CH230-20A8, WORKING DRAFT SEQUENCE.
 ACCESSION AC103481 GI:30520434
 VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Rattus norvegicus
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Ayalew, D., Ayalew, D., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismail, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Cavazos, I., Caesar, H., Center, A., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Dreger, H., Dusan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Garbargovits, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Louie, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Ntambi, C., Ntambi, C., Okunolu, G., Olamugbo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., L., Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, U., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Vales, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Wotley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, U., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
 JOURNAL
 Unpublished

REFERENCE
 AUTHORS Worley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
 On May 10, 2003 this sequence version replaced gi:25092234. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GHSS
 Center clone name: CH230-20A8
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 209990 bases at least Q40
 Consensus quality: 212566 bases at least Q30
 Consensus quality: 214461 bases at least Q20
 Estimated insert size: 223285; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Restrained insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence.
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 224302: contig of 224302 bp in length.
 location/Qualifiers
 1. 224302
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-20A8"
 /complement(224302..223315)
 /note="clone_boundary
 clone_end:77
 site:BCORI
 end_sequence:BH317689"

ORIGIN
 Query Match 62.9%; Score 25.8; DB 14; Length 224302;
 Best Local Similarity 81.1%; Pred. No. 95;
 Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

2 TTAGGAGTCATATGAAAAACAATATTTGGGTATAC 38
 75327 TTAGGAGTCATATGAAAAACAATATTTGGGTATAC 75363

RESULT 8
LOCUS AR271714 41 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 32 from patent US 6503744.
ACCESSION AR271714
VERSION AR271714.1 GI:29703259
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent:US 6503744-A 32-07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

Db 1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

RESULT 9
LOCUS AR481803 41 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 46 from patent US 6699705.
ACCESSION AR481803
VERSION AR481803.1 GI:47243438
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent:US 6699705-A 46-02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

Db 1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

RESULT 10
LOCUS AR527402 41 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 46 from patent US 6723545.
ACCESSION AR527402
VERSION AR527402.1 GI:53914423
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent:US 6723545-A 46-20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

Db 1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

RESULT 11
LOCUS AR609681 41 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 46 from patent US 6825019.
ACCESSION AR609681
VERSION AR609681.1 GI:5664981
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent:US 6825019-A 46-30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

Db 1 CTTAGAGGTCATATGAAAAAGTTATTATTGCTG 35

RESULT 12
LOCUS AX934469 41 bp DNA linear PAT 05-JAN-2004
DEFINITION Sequence 46 from Patent WO2074942.
ACCESSION AX934469
VERSION AX934469.1 GI:40641707
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent:WO 02074942-A 46-26-SEP-2002;
National Research Council of Canada (CA)
FEATURES
source Location/Qualifiers
1..41

ORIGIN

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence:CU-131 5' primer
used to amplify and clone ORF 7a"

Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7,8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Cy 1 CTTAGAGGTGATGCAAAACAAATATGCGG 35
1 CTTAGAGGTGATGCAAAAGTATATTCTG 35

RESULT 13
AE015576 10029 bp DNA linear BCT 02-DEC-2002
LOCUS Shewanella oneidensis MR-1 section 125 of 457 of the complete
DEFINITION genome
ACCESSION AE015576 AE014299
VERSION AE015576.1 GI:24347020
KEYWORDS
SOURCE
ORGANISM

Shewanella oneidensis MR-1
Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
Shewanellaceae; Shewanella.

REFERENCE
AUTHORS
1 (bases 1 to 10029)
Heidelberg, J., Paulsen, I., Nelson, K., Gaidos, E., Nelson, W.,
Read, T., Eisen, J., Seshadri, R., Ward, N., Meche, B., Clayton, R.,
Meyer, T., Tsaplin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S.,
Deboy, R., Dodson, R., Durkin, A., Haft, D., Kolonay, J., Madupu, R.,
Peterson, J., Umayam, L., White, O., Wolf, A., Yamahayan, J.,
Weidman, J., Impraim, M., Lee, K., Berry, K., Lee, C., Mueller, J.,
Khoury, H., Gill, J., Utecherback, T., McDonald, L., Feldblyum, T.,
Smith, H., Venter, J., Nealeon, K., and Fraser, C.

Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis

Nat. Biotechnol. 20 (11), 1118-1123 (2002)
12368813

2 (bases 1 to 10029)
Heidelberg, J., Paulsen, I., Nelson, K., Gaidos, E., Nelson, W.,
Read, T., Eisen, J., Seshadri, R., Ward, N., Meche, B., Clayton, R.,
Meyer, T., Tsaplin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S.,
Deboy, R., Dodson, R., Durkin, A., Haft, D., Kolonay, J., Madupu, R.,
Peterson, J., Umayam, L., White, O., Wolf, A., Yamahayan, J.,
Weidman, J., Impraim, M., Lee, K., Berry, K., Lee, C., Mueller, J.,
Khoury, H., Gill, J., Utecherback, T., McDonald, L., Feldblyum, T.,
Smith, H., Venter, J., Nealeon, K., and Fraser, C.

Direct Submission
Submitted (12-SEP-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

FEATURES
source
1. 10029
/organism="Shewanella oneidensis MR-1"
/mol_type="genomic DNA"
/strain="MR-1"
/db_xref="taxon:211586"
complement (93..584)
/locus_tag="SO1318"
complement (93..584)
/locus_tag="SO1318"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAN54383.1"
/db_xref="GI:24347021"
/translation="MRILRLVCLIFVPALFPAPHVANGEMPARSOQLAVKYMALTE
HDYQTSKRYNDSDVPEDKTARKTKGRIIDPFPRABEGVLEVDYDTEHHYNNAGSL
VMIGNYHFKGEBOFGKPKITDAIPVSTLKLDMKTRHVAEHMDLIDYQTLSDQL
AMQ"

gene

CDS

complement (586..1212)
/locus_tag="SO1319"
complement (586..1212)
/locus_tag="SO1319"
/note="similar to GP:6606500; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAN54384.1"
/db_xref="GI:24347022"
/translation="MNMWTFPFDLCGTAVFALSGALAGRHPDFGVIVASVAVG
GGSRLDILGATPVFVIRDPYIVIVATVACLLVRRPKMEYVLPVADALAL
FTVIGARALMNGISGMIAVMGILTVGGGIIIDLCROIPLVLRTEIVYTAIIIG
IGTVSLACGKERTALPLAMASALIRLCAIKHLSIPADLTKE"

gene

CDS

complement (1209..1598)
/locus_tag="SO1320"
complement (1209..1598)
/locus_tag="SO1320"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAN54385.1"
/db_xref="GI:24347023"
/translation="MFMNMYSLTTPALLFPALISLLLAATNRPALALIRIOLSGD
KPVQDQKNSQRIIRIRRMQVAGVSFALCVLCMTFIYIGLAKTGSITFGASLLLL
LSLILSVETIRISVDALNHIKEMSK"

gene

CDS

complement (1642..2631)
/locus_tag="SO1321"
complement (1642..2631)
/locus_tag="SO1321"
/note="similar to GB:M73980, SP:P46531, and PID:338675;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="cobd-related protein"
/protein_id="AAN54386.1"
/db_xref="GI:24347024"

gene

CDS

complement (2663..3373)
/gene="pfs"
/locus_tag="SO1322"
complement (2663..3373)
/gene="pfs"
/locus_tag="SO1322"
/note="similar to GB:X70848, GB:J04759, GB:M63960,
SP:P08129, SP:P36873, SP:P37140, PID:190281, PID:190516,
PID:35451, PID:484316, PID:506772, and PID:531476;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="5-methylthioadenosine
nucleosidase/S-adenosylhomocysteine nucleosidase"
/protein_id="AAN54387.1"
/db_xref="GI:24347025"

gene

CDS

3372..3578
/locus_tag="SO1323"
3372..3578
/locus_tag="SO1323"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=1

```
/product="hypothetical protein"
/protein_id="AA054388.1"
/db_xref="GI:24347026"
/translation="MGDDPDSCTCVPLMMMDGMAANIPOLCFGVIVGLRLDRDNT
KQGLGKNAHLKRAAFYSSFVISHL"
complement(3641..5047)
/gene="gltD"
/locus_tag="SO1324"
complement(3641..5047)
/gene="gltD"
/locus_tag="SO1324"
/notice="similar to SP:P09832; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamate synthase, small subunit"
/protein_id="AA054389.1"
/db_xref="GI:24347027"
/translation="MSNDPQFLTEVGRKDPTEKPAARSTQPIETIYQPAQVSEQAD
RLDGNPYCEWKCPLAHYIPMLKLAQOGRIMEADLVHENTNPEIGRCPODRL
CEGATLMDPGAVTIGVETKYITDIAOGRRPMKSKVPRKXVATVAGAPAGLC
ADILARNGKAVPDKYPOIGELITGIPAKLDKAVATRTVTEGMTEKLGTV
GKDAFNPQLRETDVFLGWTYPAKAGLEBDQGVYQALPYLIGTHLMGTSIP
DMPYTLNAGKRVVILGGEDTAMD CVRTAVROGASVTCAYRDEANMPSRREVOAR
BEGVNFLENRQVVAIKTODGKTVGECVETQMGKADASGRORAEAVIGSEOLLADAV
IAPGPOSPAPMPADYGIELDPWGKRVKASKQANDPFTTNKVFAGGDMVAGSLVV
TAIARGRDAAGQILNVLDD"
complement(5054..9502)
/gene="gltB"
/locus_tag="SO1325"
complement(5054..9502)
/gene="gltB"
/locus_tag="SO1325"
/notice="identified by match to PPM protein family HMM
PF00310"
/codon_start=1
/transl_table=11
/product="glutamate synthase, large subunit"
/protein_id="AA054390.1"
/db_xref="GI:24347028"
/translation="MSLYHPSFERDNCGFLIADNDGEASHRIYRTAIGLDRMKRG
GIASDRTGDCGLMOLPLQFEVAVAEENDMLSRKAVGMLFLSODEBLDAQFI
LERELRETLISAGRKVPVNDVLSIGKSLPIQVLYLAPIGMRKEDLERLYM
ARRLEOOLITDDKRVYVLSGQVIVYKGLMPADLPAPYPLADIRLKSSTICPHOR
PSTNTPSPKPLAOPRYLAHNEITITGNROMAARAKRNSPLIPLOQAAPVNE
TGSBSSIDNMLBMLSGMDIRARMLIIPAWOSNPEMDLCAFYDPSMPEPW
DGPAGIWMINGHAAQAVDRNGLRSPRYVITGDRILTLASEGMDVADYDEVTEKRY
GPELLIVDTLNGRLYOSFEIDNDRKHPYKEMAKNSRTLIPALQTLRQHGSEL
SPBOLYOXOKOBYTRELEQVITVLAKEGEEATGSMGDDPMAVLSKRSLLYVFR
OKFAVNTPIDPLAEKKVMSLATCIGRONFNTTGHAYRVWNSPILILFSDNOL
LGDSYTRANI VDLNTPKESGLEAIRITSEARLARSGTLLILSDRAIDKSAQV
IPVANAQVQOVLVSKSLRCDNTNII VETASARDHFAVLFGATAIYPLVYESI
SLCAKGRDLDITLALMLNFRYGIKGLKIMSKMGIISTIGSYRCQCFEALGASDI
ELCFKGVISRIEGASPEDIANQALHKAARYAHLPOGGLKTYDGEVHCFEPDV
VNTLOASLIDKNFATYKKKFAELVDRNPATLRDLGIRAGOTAIELANNVHAASNTYPR
PDSAMSTGALSPEAHALATAMNLDGRRSNSGEGEDARPRNTERNSAIQIASARF
GVTHAYVYNAVYLOIKVAQAKPGEQGLPEKRSVETAGLHARPGVTLISPEPHD
IYISIEDLAQILFDELQINTKALISVLAIVSEBQVGIATGAAKAYADMTLISGYDGTG
ASPIVSYKASPMELGLAEVHOSLVENGLRHKRILOVDGGLKTGTDVITKALLGAEF
FGFTVPMIALGCKYLRI CHLNNCATGATODKLDNDHNGLPERWVTEFVAYEV
REMAATLGVSKPEDLVGRSEWMLTLEGGTDORGLDAPILYOPNVRASRTQWETN
PPADILNHOHLNECOSAVDKGSPDAYSINNTDRVGARLSGITATTGIVKSKA
PIKSPKSAQSGRQVWNSPGIBELKCGDANDYVKGMSGGCTIVYPLVSPPOSERS
AIVGNTCYGATGGRFPAAQAGERFAVRNSGALA VVGLDNGCEVNTSGI VVVLGK
TGVMFGAOTGFAVVDQFGRFNRRVNSELVDTQKSLPQQOHLKGLIREHVALETG
SEHAMMILSDPANMIDCEVLVYKPIIAVADLLKLEQSSPELAVKAG"
```

```
Query Match 61.5%; Score 25.2; DB 1; Length 10029;
Best Local Similarity 78.9%; Pred. No. 2.8e+02;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
3 TAGAGGTCAATGAGAAAAACAAATATTCGGGTATA 40
```

```
QY
```

```
Db 388 TTGGCGGCTTATCGAAAAAACACTATCGCGGTATA 425
RESULT 14
AE015943/c 301439 bp DNA linear BCT 06-FEB-2003
LOCUS Clostridium tetani E88, section 8 of 10 of the complete genome.
DEFINITION AE015943 AE015927
ACCESSION AE015943.1 GI:28204047
VERSION AE015943.1
KEYWORDS
SOURCE Clostridium tetani E88
ORGANISM Clostridium tetani E88
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
1 (bases 1 to 301439)
Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
Gottschalk, G.
The genome sequence of Clostridium tetani, the causative agent of
tetanus disease
Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
2 (bases 1 to 301439)
Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
Gottschalk, G.
Direct Submision
Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
of Microbiology and Genetics, Georg-August University,
Grisebachstr. 8, Goettingen 37077, Germany
Location/Qualifiers
1..301439
/organism="Clostridium tetani E88"
/mol_type="genomic DNA"
/strain="Massachusetts"
/sub_strain="E88"
/db_xref="taxon:212717"
92..778
/locus_tag="CTC01981"
92..778
/locus_tag="CTC01981"
/codon_start=1
/transl_table=11
/product="L-serine dehydratase beta subunit"
/protein_id="AA036487.1"
/db_xref="GI:28204048"
779..1651
/locus_tag="CTC01982"
779..1651
/locus_tag="CTC01982"
/codon_start=1
/transl_table=11
/product="L-serine dehydratase alpha subunit"
/protein_id="AA036488.1"
/db_xref="GI:28204049"
/transl_table=11
/translation="MFVNTGLBLIDICKENKLIKWEGLNVELEBETGETKDFLIDKN
KTLVQKQSANGLKEKYSVGLIGDVAAYGANTYSBDKNTLTSGLVYKAAARALSSS
ENNAAMKIVQIPPTAGSGGILIPALISAGEALNSDEDLVALFTASVGVYIAKNAT
LAGEGGCAQACGSAAMAASIVELMGCTVDSGLDAAYIVIKNIIGLVCPVAGIVE
IPCKRVAAGVVSALTSAADNMAVSKAIPDDIVLAMYKVRQLPHELRATALGGLA
TTPAGLTKKRVK"
```

```
gene 11704..2639
complement(11704..2639)
CDS 11704..2639
complement(11704..2639)
/locus_tag="CTC01983"
/locus_tag="CTC01983"
/codon_start=1
/transl_table=11
```

gene
2788..3168
/gene="gntR"
/locus_tag="CTC01984"
2788..3168
/gene="gntR"
/locus_tag="CTC01984"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, gntR family"
*
/protein_id="AAO36490.1"
/db_xref="GI:28204051"
/translation="MINIDSSRSKSIYQOIVDOIKENILKGLKPGDKLPSYEMSL
LTVNNTYKAVQELEROKTETJRGRTYYSRNYKRRMDDKLFIKEELKLLIEA
LYMGDKKFINILDELKDFERM"
3173..4069
/locus_tag="CTC01985"
3173..4069
/locus_tag="CTC01985"
/codon_start=1
/transl_table=11
/product="ABC transporter ATP-binding protein"
/protein_id="AAO36491.1"
/db_xref="GI:28204052"
/translation="MLEISNVNKKLGNLLNNVNLVNSKSIPIGLIGENAGAKTLLI
KCIITVYDNGENIKIFNBEVFNENTKIKKGIYADENQYIPSPVSEMLKPFGLTYK
KESMRNFEINLNTKPIPIDKRIKELSKGMKRLSLMNLGIMPEILLDPSGLDPLV
IKKVIYKNDVQAKNNNTTIFISHHIGDLERICDNVAFINKGITSYNNINDMKHI
KKQVFPKNDPDKIQSDMEVLTIEISGRVNYVITKFSKLEELKMSMALPIEELD
LSLEMFIVSYSEEDHNEKILA"
4041..5114
/locus_tag="CTC01986"
4041..5114
/locus_tag="CTC01986"
/codon_start=1
/transl_table=11
/product="membrane spanning protein"
/protein_id="AAO36492.1"
/db_xref="GI:28204053"
/translation="NNTMKKPLFNDPLFYKDFPKSARMTSLILLVLYIKRLLIPSPILN
TIKSNLTPTQISPGNMTPKRAEMAYVWRKNLLASDQDMVITITFVILLIVLKKER
NNSTYSFPRASMPKREBEIVKIKLAGNSISIPPIPTILLISFYKNTKWINDPYII
VQWITINILFTALFSFIFPIQTMGQNTIYASITGALVGIPEACIVMSITDILTSY
KNPIIKINILFTALFSFIFPIQTMGQNTIYASITGALVGIPEACIVMSITDILTSY
VSLTGLFPLASVYPTDKCIKYYSK"
5132..5824
/locus_tag="CTC01987"
5132..5824
/locus_tag="CTC01987"
/codon_start=1
/transl_table=11
/product="transporter"
/protein_id="AAO36493.1"
/db_xref="GI:28204054"
/translation="MDYITLBCNDLVKKGKTIYANNLTKIPKKGITGLGPNQVGT
TILKIVYASLNTNSQITITDGSPGVDTIKIYVAPLDEEYFQWMSIAEALNPKFSF
EEDPVKVASLNTNLNLSKOKIKSLSGMQEKNTISLTFSRKSLTIDEEPLAAADV
VRDRIIEIILNPDKNSSIIISTHLLINDIYLFNEVVFINDGKTIILGHVLNKKYK
KKSIDLFKEIRI"
5825..6607
/locus_tag="CTC01988"
5825..6607
/locus_tag="CTC01988"
/codon_start=1

[illegible]

Search completed: April 7, 2006, 18:04:55
 Job time : 285.271 secs

```

DEFINITION      Human DNA sequence from clone RP11-44D23 on chromosome 10, complete
                  sequence.
ACCESSION       AL589943
VERSION         AL589943.26  GI:20068452
KEYWORDS        HTG.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                 Homiidae; Homo.
REFERENCE       1  (bases 1 to 139325)
AUTHORS         Pelan,S.
TITLE           Direct Submission
JOURNAL         Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
                  Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
                  Clone requests: clonerequest@sanger.ac.uk
                  On Apr 7, 2002 this sequence version replaced gi:19848045.
                  The following abbreviations are used to associate primary accession
                  numbers given in the feature table with their source databases:
                  EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
                  on the WORMPEP database can be found at
                  http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
                  was generated from part of bacterial clone contigs of human
                  chromosome 10, constructed by the Sanger Centre Chromosome 10
                  Mapping Group. Further information can be found at
                  http://www.sanger.ac.uk/HGP/Chr10
                  RP11-44D23 is from the library RPCT-11.1 constructed by the group
                  of Pieter de Jong. For further details see
                  http://www.chori.org/bacpac/home.htm
                  VECTOR: pBACe3.6

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

Location/Qualifiers
1..139325
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="10"
  /clone="RP11-44D23"
  /clone_1fb="RPCT-11.1"
  /clone_2fb="RPCT-11.1"
  /note="Clone_right_end: RP11-49L2"
  /note="Tandem repeat. Forced join. Gap size estimated to
be approximately 170bp by restriction digest data."
  /note="Single clone region. Assembly confirmed by
restriction digest data."
  /note="Clone_left_end: RP11-265G8"
  /note="Clone_right_end: RP11-44D23"
  /note="Clone_right_end: RP11-44D23"

ORIGIN
Query Match      60.5%; Score 24.8; DB 8; Length 139325;
Best Local Similarity 80.6%; Pred. No. 2.2e+02;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 AGAGGTCAATATGGAACAAATATTTGCGTTAT 39
DB 118620 AGGATGCATATGAGAAAGAAATATTCCTTTAT 118585

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 1 7, 2006, 15:58:06 ; Search time 46.2484 Seconds
(without alignments)
5908.370 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 cttagagagctcatatcgtaa.....acaaatattcggtttac 41

Scoring table:

IDENTITY_NTC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	41	2	AAx84282
2	29	70.7	65632	3	AAAB1502
3	29	70.7	110000	3	AAAB1490_00
4	29	70.7	349980	3	AAFP1544
5	28	68.3	687	3	AAZ24472
6	25.4	62.0	41	3	AAAS3735
7	25.4	62.0	41	3	ABT13687
8	24.8	60.5	165199	6	ABR83460
9	24.4	59.5	4356	12	AD020395
10	24.4	59.5	7183	11	ACN90426
11	23	56.1	444	3	AAC97275
12	23	56.1	1024	6	ABX65592
13	23	56.1	1680	8	ACA34620
14	23	56.1	1780	2	AAAI4256
15	23	56.1	4391	14	AAAB1394
16	22.6	55.1	1538	6	ABQ70554
17	22.6	55.1	1628	3	AAC47551
18	22.6	55.1	1633	3	AAC35365
19	22.6	55.1	2000	8	ADA69010

C 20	22.6	55.1	14012	8	AA050004	AA050004 Human Ras
C 21	22.6	55.1	110000	6	ABA03041_19	Continuation (20 C
C 22	22.4	54.6	542	4	AAK62272	AAK62272 Human imm
C 23	22.4	54.6	1963	3	AAK83550	AAK83550 DNA encod
C 24	22.4	54.6	4121	6	ABL59256	ABL59256 Adenosine
C 25	22.4	54.6	4954	4	AAK82088	AAK82088 Human imm
C 26	22.2	54.1	3089	4	ABL26828	ABL26828 Drosophila
C 27	22.2	54.1	4936	6	ABJ32111	ABJ32111 Human imm
C 28	22	53.7	31	3	AAZ24470	AAZ24470 N. mening
C 29	22	53.7	62	2	AAV81717	AAV81717 Neisseria
C 30	22	53.7	690	8	ACA29466	ACA29466 Prokaryot
C 31	22	53.7	1263	8	ACA28653	ACA28653 Prokaryot
C 32	22	53.7	1401	9	ADA32829	ADA32829 DNA encod
C 33	22	53.7	1422	10	ADC92660	ADC92660 B. faeciu
C 34	21.8	53.2	550	5	ABV56670	ABV56670 Human pro
C 35	21.8	53.2	1734	13	ADT46561	ADT46561 Bacterial
C 36	21.8	53.2	6075	2	AAQ27451	AAQ27451 Type A hu
C 37	21.8	53.2	25944	11	ADL27125	ADL27125 Mouse gen
C 38	21.8	53.2	26317	9	ADA66331	ADA66331 Mouse IgG
C 39	21.8	53.2	26320	9	ADA03047	ADA03047 Mouse IgG
C 40	21.8	53.2	26320	10	ADB72785	ADB72785 Mouse IgG
C 41	21.8	53.2	74828	12	ADQ97992	ADQ97992 Human can
C 42	21.8	53.2	110000	14	ADZ13747_1	Continuation (2 of
C 43	21.8	53.2	347814	12	ADQ59440	ADQ59440 Human can
C 44	21.4	52.2	1020	10	ACF68472	ACF68472 Phototrab
C 45	21.4	52.2	2838	10	AD69742	AD69742 Bacterial

ALIGNMENTS

RESULT 1	AAx84282
ID	AAx84282 standard; DNA; 41 BP.
XX	AAx84282;
AC	XX
DT	08-SEP-1999 (first entry)
XX	XX
DE	PCR primer for Neisseria CMP-Neu5Ac synthetase coding sequence.
XX	XX
KW	Beta-1,4-galactosyltransferase; lgtB; fusion protein; catalytic domain;
KW	glycosyltransferase; accessory enzyme; nucleotide sugar formation;
KW	saccharide donor; oligosaccharide synthesis; CMP-Neu5Ac synthetase;
KW	carbohydrate structure development; PCR primer; ss.
OS	Synthetic.
OS	Neisseria sp.
XX	XX
PN	WO9931224-A2.
XX	XX
PD	24-JUN-1999.
XX	XX
PF	15-DEC-1998; 98WO-CA001180.
XX	XX
PR	15-DEC-1997; 97US-0069443P.
PR	14-DEC-1996; 96US-00211691.
XX	XX
PA	(CANADA) NAT RES COUNCIL CANADA.
PI	Gilbert M, Young NM, Wakarchuk MW;
XX	XX
DR	WPI, 1999-395174/33.
XX	XX
PT	A new glycosyltransferase fusion protein useful in the enzymatic
PT	synthesis of oligosaccharides.
XX	XX
PS	Example 1; Page 40; 63pp; English.
XX	XX
CC	This sequence represents a PCR primer for the Neisseria CMP-Neu5Ac
CC	synthetase coding sequence. The invention relates to a nucleic acid
CC	encoding a fusion protein that comprises a glycosyltransferase catalytic
CC	domain and a catalytic domain from an accessory enzyme that is involved

in formation of a nucleotide sugar which is a saccharide donor for a glycosyltransferase reaction. The fusion protein is useful in the enzymatic synthesis of oligosaccharides. The fusion proteins are able to catalyze more than one reaction involved in the enzymatic synthesis. This is useful for the development of therapeutic agents that have specific carbohydrate structures. Carbohydrates are involved in recognition elements on the surface of cells. The fusion protein can be used for the synthesis of both natural carbohydrates and synthetic derivatives with novel properties. The fusion polypeptide allows two glycosyltransferase reactions in a single vessel, provides improved yields of end products. Additionally, cleanup and disposal of extra solvents and by-products is reduced. The fusion protein can also use directly different donor analogues and various acceptors with a terminal galactose residue

Sequence 41 BP; 16 A; 5 C; 9 G; 11 T; 0 U; 0 Other;

Query Match
Best Local Similarity 100.0%; Score 41; DB 2; Length 41;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTTAGAGGTCATATGAAAAAATAATTCGCGTTATAC 41
1 CTTAGAGGTCATATGAAAAAATAATTCGCGTTATAC 41

RESULT 2
AAA81502
ID AAA81502 standard; DNA; 65632 BP.

AAA81502;
04-DEC-2000 (first entry)

N. meningitidis partial DNA sequence gmm_49 SEQ ID NO:49.

Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
antigen; vaccine; diagnosis; infection; antibacterial; identification;
Meningococcus B; MenB; ds.

Neisseria meningitidis.
WO200022430-A2.

20-APR-2000.

08-OCT-1999; 99WO-US023573.

09-OCT-1998; 98US-0103794P.

30-APR-1999; 99US-0132068P.

(CHIR) CHIRON CORP.

Frazier CM, Hickey E, Peterson J, Tettelin H, Venter JC,
Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V,
Rapponi R, Pizzo M;
WPI; 2000-318079/27.

Isolated nucleotide sequences of Neisseria meningitidis which can be used in the diagnosis and treatment of N. meningitidis infection and other Neisserial infections, for example, N.gonorrhoea.

Claim 7; Page 1331-1350; 1760pp; English.

The present invention describes methods of obtaining immunogenic proteins from Neisseria genomic sequences. AAA81453 to AAA8214 represent specifically claimed Neisseria meningitidis genomic DNA sequences. AAA81260 to AAA81303 and AA825620 to AA825663 represent Neisseria DNA sequences and their corresponding proteins; AAA81254 to AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the isolation of Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF sequences, which are all used in the exemplification of the present invention. The nucleic acid

sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to Neisserial bacteria. For example, some of the identified proteins could be components of vaccines against Meningococcus B against all serotypes; and/or against all pathogenic Neisseriae. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious Meningococcus B vaccines have failed mainly due to antigen tolerance. Multivalent CC vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may provide an opportunity to identify secreted or surface exposed proteins that may be presumed targets for the immune system and which are not antigenically variable or at least more conserved than other more variable regions

Sequence 65632 BP; 16704 A; 16327 C; 15474 G; 17136 T; 0 U; 1 Other;

Query Match
Best Local Similarity 70.7%; Score 29; DB 3; Length 65632;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

13 TATGAAAAACAAATATTCGCGTTATAC 41
55054 TATGAAAAACAAATATTCGCGTTATAC 55082

RESULT 3
AAA81490_00/C
WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490

Fragment Name	Begin	End
AAA81490_00	1	110000
AAA81490_01	100001	210000
AAA81490_02	200001	310000
AAA81490_03	300001	410000
AAA81490_04	400001	510000
AAA81490_05	500001	610000
AAA81490_06	600001	710000
AAA81490_07	700001	810000
AAA81490_08	800001	910000
AAA81490_09	900001	1010000
AAA81490_10	1000001	1110000
AAA81490_11	1100001	1210000
AAA81490_12	1200001	1310000
AAA81490_13	1300001	1410000
AAA81490_14	1400001	1437668

AAA81490 standard; DNA; 1437668 BP.

AAA81490;
04-DEC-2000 (first entry)

N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.

Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
antigen; vaccine; diagnosis; infection; antibacterial; identification;
Meningococcus B; MenB; ds.

Neisseria meningitidis.
WO200022430-A2.

20-APR-2000.

08-OCT-1999; 99WO-US023573.

09-OCT-1998; 98US-0103794P.

30-APR-1999; 99US-0132068P.

(CHIR) CHIRON CORP.

Frazier CM, Hickey E, Peterson J, Tettelin H, Venter JC;

PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
PI Rappoli R, Piza M;
XX WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be used
PT in the diagnosis and treatment of *N. meningitidis* infection and other
PT *Neisseria* infections, for example, *N. gonorrhoea*.
XX
PS Claim 7; Page 866-1272; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from *Neisseria* genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed *Neisseria meningitidis* genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent *Neisseria* DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC *Neisseria meningitidis* DNA sequences; and AAA81322 to AAA81452 represent
CC *Neisseria meningitidis* MemB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to *Neisseria* bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against *Neisseria*. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious *Meningococcus B*
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC antigens have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 1437668 BP; 344338A; 353206C; 385074G; 355045T; 0U; 50ther;
XX
Query Match 70.7%; Score 29; DB 3; Length 110000;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 13 TATGAAAAAACAATATTGCGGTATAC 41
DB 79104 TATGAAAAAACAATATTGCGGTATAC 79076
XX
RESULT 4
AAF21544/c
ID AAF21544 standard; DNA; 349980 BP.
XX
AC AAF21544;
XX
DT 13-MAR-2001 (first entry)
XX
DE *Neisseria meningitidis* B nucleotide sequence SEQ ID NO.1.
XX
KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS *Neisseria meningitidis*.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000. 4.
XX
PF 08-MAR-2000; 2000WO-US005928.
XX
PR 30-APR-1999; 99US-013206BP.
PR 08-OCT-1999; 99WO-US023573.
PR 28-FEB-2000; 2000GB-00004695.

XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Piza M, Hickey B, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
PI Rappoli R, Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT *Neisseria meningitidis* B full length genome sequence and open reading
PT frames are used to detect, treat and prevent *Neisseria* infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of *Neisseria*
CC *meningitidis* B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC AAF21608, and so on). AAF21545 to AAF21588 encode the *Neisseria* proteins
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. *Neisseria* nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to *Neisseria* bacteria or as a
CC diagnostic reagent for detecting the presence of *Neisseria* bacteria or
CC of antibodies raised to *Neisseria* bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used
XX
SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 U; 0 Other;
XX
Query Match 70.7%; Score 29; DB 3; Length 349980;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 13 TATGAAAAAACAATATTGCGGTATAC 41
DB 79104 TATGAAAAAACAATATTGCGGTATAC 79076
XX
RESULT 5
AAZ24472
ID AAZ24472 standard; DNA; 687 BP.
XX
AC AAZ24472;
XX
DT 17-FEB-2000 (first entry)
XX
DE *N. meningitidis* slab DNA.
XX
KW CMP sialate synthetase; slab; cytidine monophosphate; detection;
KW CMP-activated N-acetylneuraminic acid; sialyl acceptor; ds.
XX
OS *Neisseria meningitidis*.
XX
PN DE19913206-A1.
XX
PD 07-OCT-1999.
XX
PF 24-MAR-1999; 99DE-01013206.
XX
PR 26-MAR-1998; 98DE-01013426.
XX
PA (FESS/) FESSNER W D.
XX

PI Resener W, Knorret M;
 XX WPI; 2000-024923/03.
 DR P-PSDB; AAY50798.
 XX
 PT Production of CMP-sialate synthetase protein useful for producing CMP-
 PT activated N-acetylneuraminic acid analogs and non-naturally sialylated
 PT oligosaccharides and glycoconjugates.
 XX
 PS Claim 5; Page 19-20; 26pp; German.
 XX
 CC This invention describes a novel method for the production of a CMP
 CC (cytidine monophosphate)-sialate synthetase protein which comprises
 CC culturing a prokaryotic host organism transformed with an expression
 CC vector containing a structural gene for the protein, where the start
 CC codon of the structural gene is 8-12 bases downstream from a ribosome
 CC binding site (RBS). The protein is useful for producing CMP-activated N-
 CC acetylneuraminic acid analogs by reacting the analogs with cytidine
 CC triphosphate in the presence of the protein and for producing non-
 CC naturally sialylated oligosaccharides and glycoconjugates, which are of
 CC therapeutic interest because of their sialidase resistance. The protein
 CC is also useful for detecting sialyl acceptors by reacting the acceptor
 CC with a fluorescent sialic acid derivative in the presence of the protein,
 CC cytidine triphosphate and a sialyl transferase and detecting the
 CC fluorescently modified acceptor by selective chromatography. This
 CC sequence encodes a CMP (cytidine monophosphate)-sialate synthetase
 CC protein, sIab, isolated from *Neisseria meningitidis* serogroup B
 SQ Sequence 687 BP; 218 A; 154 C; 122 G; 193 T; 0 U; 0 Other;

Query Match 68.3%; Score 28; DB 3; Length 687;
 Best Local Similarity 100.0%; Pred No. 1.7; Mismatches 0; Indels 0; Gaps 0;

OY 14 ATGGAACCAAAATTTGCGGTTATAC 41
 |||||
 DB 1 ATGGAACCAAAATTTGCGGTTATAC 28

RESULT 6
 AAA53735
 ID AAA53735 standard; DNA; 41 BP.
 XX
 AC AAA53735;
 XX
 DT 22-DEC-2000 (first entry)
 XX

DE Primer for amplifying ORF 7a from LPS core biosynthesis locus.
 XX
 KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
 KW immunity; immunogen; ganglioside; primer; ss.
 XX
 OS Campylobacter jejuni.
 XX
 PN WO200046379-A1.
 XX
 PD 10-AUG-2000.
 XX
 PF 01-FEB-2000; 2000MO-CA000086.
 XX
 PR 01-FEB-1999; 99US-0118213P.
 XX
 PR 31-JAN-2000; 2000US-00495406.
 XX
 PA (CANA) NAT RES COUNCIL CANADA.
 XX
 PI Gilbert M, Wakarchuk WW;
 XX
 DR WPI; 2000-524418/47.
 XX
 PT Novel glycosyltransferase polypeptides and polynucleotides useful for

PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
 PT reagents and as immunogen for producing antibodies.
 XX
 PS Example 1; Page 47; 120pp; English.
 XX

CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
 CC useful for synthesizing sialylated oligosaccharide such as ganglioside,
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
 CC other oligosaccharides that have biological activity. The enzymes and
 CC nucleic acids that encode them are useful for studies of the pathogenesis
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
 CC jejuni and the nucleic acids are used as probes to study expression of
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analyzing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing
 CC reagents or as therapeutics and as immunogens for producing antibodies.
 CC Bacterial glycosyltransferase can be used to catalyse the formation of
 CC oligosaccharides that are identical to the corresponding mammalian
 CC structures and are easier and less expensive to produce in large
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial
 CC origin of the enzymes facilitates expression of large quantities of the
 CC enzymes using relatively inexpensive prokaryotic expression systems. Two
 CC primers (AAA53735, AAA53736) were used to amplify open reading frame 7a
 CC from the LPS core biosynthetic locus
 SQ Sequence 41 BP; 15 A; 3 C; 10 G; 13 T; 0 U; 0 Other;

Query Match 62.0%; Score 25.4; DB 3; Length 41;
 Best Local Similarity 82.9%; Pred No. 11; Mismatches 6; Indels 0; Gaps 0;

OY 1 CTTAGAGGTCATATGAAAAACAAATATTGCGG 35
 |||||
 DB 1 CTTAGAGGTCATATGAAAAACAAATATTATTGCTG 35

RESULT 7
 ABL13687
 ID ABL13687 standard; DNA; 41 BP.
 XX
 AC ABL13687;
 XX
 DT 07-FEB-2003 (first entry)
 XX

DE Campylobacter jejuni bifunctional sialtransferase cslII PCR primer #3.
 XX
 KW PCR; primer; enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylglucosamine transferase; ss;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.
 XX
 OS Campylobacter jejuni.
 XX
 PN WO200274942-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002MO-CA000229.
 XX
 PR 21-MAR-2001; 2001US-00816028.
 XX
 PR (CANA) NAT RES COUNCIL CANADA.
 XX
 PA Gilbert M, Wakarchuk WW;
 XX
 DR WPI; 2003-040554/03.

XX New glycosyltransferases from *Campylobacter*, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.

PS Example, Page 49, 107pp, English.

XX The invention comprises the amino acid and coding sequences of
CC *Campylobacter jejuni* proteins. The C. jejuni proteins of the invention
CC may be either an acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylglucosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC biacyl acid synthase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of *Campylobacter* enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumor
CC metastasis). The present DNA sequence represents a *Campylobacter jejuni*
CC PCR primer that was used in an example of the invention

SO Sequence 41 BP, 15 A, 3 C, 10 G, 13 T, 0 U, 0 Other;

Query Match 62.0%; Score 25.4; DB 8; Length 41;
Best Local Similarity 82.9%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATAATTTGGCG 35
DB 1 CTTAGAGGTCATATGAAAAACAATAATTTGGCTG 35

RESULT 8
ABK83460

ID ABK83460 standard; cDNA, 165199 BP.

AC ABK83460;

DT 14-AUG-2002 (first entry)

DE Human CDNA differentially expressed in granulocytic cells #31.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

OS WO200228999-A2.

PN 11-APR-2002.

PD 03-OCT-2001; 2001WO-US030821.

PF 03-OCT-2000; 2000US-0237189P.

PR (GENE-) GENE LOGIC INC.

PA Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;

PI WPI; 2002-435328/46.

DR Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX Claim 1; SEQ ID NO 31, 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the

CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammatory (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease, also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 165199 BP; 48510 A; 33223 C; 34406 G; 49060 T; 0 U; 0 Other;

Query Match 60.5%; Score 24.8; DB 6; Length 165199;
Best Local Similarity 80.6%; Pred. No. 47;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGAAAAACAATAATTTGGCGTT 37
DB 113374 TTAGAGGTCATATGAAAAACAATAATTTGGCTTT 113409

RESULT 9
ADO20395

ID ADO20395 standard; cDNA, 4356 BP.

AC ADO20395;

DT 12-AUG-2004 (first entry)

DE Human PRO polynucleotide #643.

XX Human; PRO; gene; ss; immune related disorder;
KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW vasculitis; sarcoidosis; autoimmune haemolytic anemia;
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
KW peripheral nervous system; demyelinating disease; central nervous system;
KW Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.

XX Homo sapiens.

OS WO2004043361-A2.

PN 27-MAY-2004.

XX	06-NOV-2003; 2003W0-US0355268.
PF	
PR	08-NOV-2002; 2002US-0425235P.
XX	
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI	Wood WJ, Wu TD;
XX	
DR	WPI; 2004-420067/39.
DR	P-PSDB; ADO20396.
XX	
PT	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT	treating an immune related disorder such as systemic lupus erythematosus,
PT	rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT	spondyloarthropathy.
XX	
PS	Claim 1; SEQ ID NO 1378; 1731bp; English.
XX	
CC	The invention relates to human PRO polypeptides and the polynucleotides
CC	encoding them. The polypeptides and polynucleotides are useful for
CC	treating and diagnosing immune related disorders in mammals. The immune
CC	related disorders include systemic lupus erythematosus, rheumatoid
CC	arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC	sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC	haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC	mellitus, immune-mediated renal disease, demyelinating diseases of the
CC	central or peripheral nervous system, demyelinating polyneuropathy,
CC	Gullain-Barre syndrome and chronic inflammatory demyelinating
CC	polyneuropathy. This sequence represents a human PRO polynucleotide of
CC	the invention.
XX	
SO	Sequence 4356 BP; 1230 A; 987 C; 1027 G; 1112 T; 0 U; 0 Other;
	Query Match 59.5%; Score 24.4; DB 12; Length 4356;
	Best Local Similarity 82.4%; Pred. No. 44;
	Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0
OY	2 TTAGAGGTCATATGAAAAACAAATATTTGCGG 35
DB	650 TTAGGAGGTCATATGAAAAATCAAGCATTTGCTG 683
RESULT 10	
ACN90426	
ID	ACN90426 standard; DNA; 7183 BP.
XX	
AC	ACN90426;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Breast cancer related marker, seq id 11576.
XX	
KM	Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX	
OS	Homo sapiens.
XX	
PN	US2003099974-A1.
XX	
PD	29-MAY-2003.
XX	
PF	18-JUL-2002; 2002US-00198846.
XX	
PR	18-JUL-2001; 2001US-0306220P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Lillie J, Xu Y, Wang Y, Steinmann K;
XX	
DR	WPI; 2003-787014/74.
PT	Novel isolated polypeptide associated with breast cancer, useful for

FT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
PS Disclosure; SEQ ID NO 11576; 36pp; English.

CC The invention relates to an isolated polypeptide (I) associated with a
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for the
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN97851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=2003099974

XX Sequence 7183 BP; 1894 A; 1719 C; 1807 G; 1763 T; 0 U; 0 Other;
SQ

Query Match 59.5%; Score 24.4; DB 11; Length 7183;
Beef Local Similarity 82.4%; Pred. No. 46;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Dy 2 TTAGGAGGTCAATGGAAAAACAATAATTTCGC 35
617 TTAGGAGGTCAATGGAAAAACAAGACATTTCTG 650

Db

RESULT 11
AAC97275/G
ID AAC97275 standard; DNA; 444 BP.
XX AAC97275;
AC
XX
DT 23-FEB-2001 (first entry)
DE Helicobacter pylori bait polypeptide nucleotide sequence #47.
XX
XX Helicobacter pylori; two-hybrid system; protein-protein interaction;
KW antibacterial; bait polypeptide; gastric ulcer; ds.
XX
OS Helicobacter pylori.
XX
PN MO200066722-A1.
PD
XX 09-NOV-2000.
PF 14-APR-2000; 2000MO-IB000603.
PR 30-APR-1999; 99EP-00401066.
PA (HYBR-) HYBRIGENICS SA.
PI Legrain P, Selig L, Rain J;
PI
XX WPI; 2000-687535/67.
DR P-PDB; AAB52529.
XX
XX A two-hybrid system for identifying compounds useful in the treatment of
PT e.g. gastric ulcers comprises producing a collection of recombinant cell
PT clones.
PT
XX
XX Example 5; Page 139; 267pp; English.
PS
XX
CC The present sequence encodes a bait polypeptide used in a Helicobacter
CC pylori two-hybrid screen to identify protein-protein interactions. The
CC method is used to identify a recombinant cell clone expressing a prey
CC polypeptide which is capable of interacting with the bait polypeptide.
CC The two hybrid system is useful for screening compounds for antibacterial
CC activity. It may be used in the treatment of gastric ulcers. The
CC polynucleotides are useful as amplification primers or specific detection
CC probes. The polypeptides, vectors or host cells can be used as immunogens

CC to produce mono- or polyclonal antibodies. The polynucleotides,
CC polypeptides, antibodies, vectors, host cells or modulating agents can be
CC used to produce a pharmaceutical composition
XX

XX Sequence 444 BP; 141 A; 75 C; 99 G; 129 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 3; Length 444;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGTAACAAATATTTGGCGTTAT 39
DB 148 CATAGTGTCTTATGCTATATAAAATCATTTGCGTTAT 110

RESULT 12

ABX65592/C
ID ABX65592 standard; DNA; 1024 BP.

XX ABX65592;

XX 07-MAY-2003 (first entry)

XX Helicobacter pylori selected interacting domain (SID) DNA #191.

XX Protein-protein interaction; ulcer; selected interacting domain; SID;
KM gene; de.

XX Helicobacter pylori.

XX WO200266501-A2.

XX 29-AUG-2002.

XX 28-DEC-2001; 2001WO-EP015428.

XX 02-JAN-2001; 2001US-0259302P.

XX (HYBR-) HYBRIGENICS.
PA (INSP) INST PASTEUR.

XX Legrain P, Rain J, Colland F, De Reuse H, Labigne A;

XX WPI; 2002-674910/72.

XX P-PSDB; ABUS0848.

XX New complexes of protein-protein interactions in Helicobacter pylori,
PT useful for identifying modulating compounds for treating or preventing
PT ulcers in mammals.

XX Claim 7; Page 132; 642pp; English.

XX The invention describes a complex of protein-protein interactions in
CC Helicobacter pylori selected from 421 complexes given in the
CC specification. The complex of protein-protein interactions are useful for
CC screening for agents which modulate the interaction of proteins.
CC Modulating compounds which binds to a targeted bacterial protein may be
CC used for treating or preventing ulcers in a human or animal. This
CC sequence encodes a selected interacting domain (SID), identified via
CC protein-protein interactions

XX Sequence 1024 BP; 344 A; 168 C; 227 G; 285 T; 0 U; 0 Other;

XX Query Match 56.1%; Score 23; DB 6; Length 1024;

XX Best Local Similarity 74.4%; Pred. No. 1.2e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGTAACAAATATTTGGCGTTAT 39
DB 150 CATAGTGTCTTATGCTATATAAAATCATTTGCGTTAT 112

RESULT 13

ACA34620/C
ID ACA34620 standard; DNA; 1680 BP.

XX ACA34620;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #16277.

XX Antisense; de; prokaryotic essential gene; cell proliferation;
KM drug design; gene.

XX Helicobacter pylori.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABUS0750.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 22490; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1680 BP; 583 A; 269 C; 339 G; 489 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 8; Length 1680;
Best Local Similarity 74.4%; Pred. No. 1.3e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CTTAGAGGTCATATGAAAAAACAATATTCGGTTAT 39
Db 1381 CATAGTGGCTTATGCTATAAAAATCATTCGTTTAT 1343

RESULT 14
AAx14256/C
ID AAx14256 standard; DNA; 1780 BP.

AC AAX14256;
DT 31-MAR-1999 (first entry)
DE H. pylori GHP0 1488 gene.
KM GHP0 protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease; ss.
XX Helicobacter pylori.
OS
FH Key Location/Qualifiers
FT CDS 51..1730
FT /*tag= a

W09843478-A1.

PD 08-OCT-1998.
PF 01-APR-1998; 98WC-US006371.

XX 01-APR-1997; 97US-00833457.
PR 24-JUN-1997; 97US-00881227.
PR 29-JUL-1997; 97US-00902615.

PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PX (HUMA-) HUMAN GENOME SCI INC.

PI Kleantous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
DR WPI; 1998-542293/46.
DR P-PSDB; AAW98537.

PT New isolated Helicobacter polymnucleotides - used to develop products for
PT the diagnosis, prevention and treatment of Helicobacter infections and
PT gastrointestinq1 diseases.

PS Claim 1, Page 1032-1034; 2054pp; English.

CC This sequence represents a polymnucleotide of the invention. It was
CC isolated from Helicobacter pylori and encodes a H.pylori GHP0 protein.
CC The polypeptides can be used for preventing or treating Helicobacter
CC infections, and gastroduodenal diseases associated with these infections,
CC including acute, chronic, and atrophic gastritis, and peptic ulcer
CC diseases, e.g. gastric and duodenal ulcers. They can also be used for the
CC production of antibodies. The products can also be used for detection and
CC diagnosis

SQ Sequence 1780 BP; 623 A; 287 C; 355 G; 515 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 2; Length 1780;
Best Local Similarity 74.4%; Pred. No. 1.3e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CTTAGAGGTCATATGAAAAAACAATATTCGGTTAT 39
Db 1431 CATAGTGGCTTATGCTATAAAAATCATTCGTTTAT 1393

RESULT 15

AEA61394
ID AEA61394 standard; DNA; 4391 BP.

AC AEA61394;

DT 11-AUG-2005 (first entry)

DE Bacillus thuringiensis Cry19Aa coding sequence, SEQ ID 4.

XX Insecticide; crystal protein; Cry19Aa; gene; ds.

OS Bacillus thuringiensis.

FH Key Location/Qualifiers
FT CDS 719..2665
FT /*tag= a
FT /product= "Cry19Aa"

PN US2005124803-A1.

PD 09-JUN-2005.

PF 30-AUG-2004; 2004US-00929754.

PR 29-AUG-2003; 2003US-0498826P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PX Dean DH, Abdullah MA;

PI WPI; 2005-417057/42.

DR P-PSDB; AEA61393.

DR EMBL; Y07603.

PT New modified Bacillus thuringiensis insecticidal crystal proteins (i.e.
PT CryIba and Cry19a) with enhanced toxicity, useful for reducing or
PT eliminating populations of target insects (i.e. mosquitoes) that are
PT vectors of disease.

PS Disclosure; SEQ ID NO 4; 63pp; English.

CC The present invention relates to modified insecticidal Bacillus
CC thuringiensis crystal proteins Cry4Ba and Cry19Aa which have enhanced
CC toxicity against a variety of insects, e.g. mosquitoes and lepidoptera.
CC The modified proteins are useful for reducing or eliminating populations
CC of target insects that are vectors of disease, particularly mosquitoes.
CC The present sequence is the coding sequence of the wild-type Cry19Aa
CC protein, which can be modified with: a substitution of amino acids at
CC positions 355 through 358 with amino acids tyrosine, glutamine, aspartic
CC acid, and leucine; an insertion of at least one amino acid (e.g.
CC arginine) after position 358 and a deletion of the amino acids at
CC positions 414 through 418.

SQ Sequence 4391 BP; 1557 A; 630 C; 914 G; 1290 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 14; Length 4391;
Best Local Similarity 74.4%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 TTAGAGGTCATATGAAAAAACAATATTCGGTTATAC 41
Db 1074 TAGAAGATCTTATAGACAAACGAATACGAGGTTATAC 1112

Search completed: April 7, 2006, 16:17:25
J0D time : 50.2484 secs

* GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:00:12 ; Search time 321.989 Seconds

(without alignments)
5957.564 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41
Sequence: 1 ctctaggagctcatatgtaaa.....acaaatattcgcgtttatc 41

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	25.8	62.9	638	9	B2245783 CH230-325
2	24.6	60.0	583	7	CR750796 DKEZp470M
3	24.6	60.0	662	11	CR869763 Sub scrof
4	24.6	60.0	701	7	CV649651 EST928404
5	24.6	60.0	771	7	CV635258 EST914011
6	24.4	59.5	500	7	CN361845 170004241
7	24.4	59.5	521	7	CN361847 170004252
8	24.4	59.5	582	3	BP350103 BP350103
9	24.4	59.5	669	1	AL703924 DKEZp686D
10	24.4	59.5	790	8	DR001464 TC114916
11	24.4	59.5	821	2	BG756622 602713774
12	24.4	59.5	905	5	BQ424531 AGENCCOURT
13	24.4	59.5	1113	3	BM557407 AGENCCOURT
14	24.4	59.5	4191	10	AY399307 Pan tlog1
15	24.4	59.5	4503	10	AY399306 Homo sabi
16	24.2	59.0	672	9	BH298686 CH230-84H
17	24.2	59.0	1190	3	BI687776 603315692
18	24.2	58.5	439	2	BI065289 pGfIn.pK0
19	24.2	58.5	531	1	AJ741063 AJ741063
20	24.2	58.5	619	5	BU296809 60373647
21	24.2	58.5	695	9	BH951067 od127411.
22	24.2	58.5	733	5	BU200026 603950538

c 23	24	58.5	773	9	BH487940
c 24	23.6	57.6	234	8	DR176547
c 25	23.6	57.6	699	2	BF681440
c 26	23.6	57.6	789	9	BZ265368
c 27	23.6	57.6	890	10	AG424476
c 28	23.4	57.1	261	1	AV164171
c 29	23.4	57.1	455	6	CF533995
c 30	23.4	57.1	546	5	BU01596 603738686
c 31	23.4	57.1	591	9	AZ382739 1M0140G08
c 32	23.4	57.1	654	6	CB521772 UI-M-GH0-
c 33	23.4	57.1	658	1	AM064320 AM064320
c 34	23.4	57.1	731	5	BU056866 UI-M-FOO-
c 35	23.4	57.1	806	6	CD350955
c 36	23.4	57.1	927	7	CJ403240
c 37	23.2	56.6	530	9	BZ443595
c 38	23.2	56.6	658	7	CJ392627
c 39	23.2	56.6	710	10	EX171933
c 40	23.2	56.6	910	10	AG893009
c 41	23.2	56.6	927	7	CJ403240
c 42	23	56.1	265	1	AJ752805
c 43	23	56.1	265	1	AJ752815
c 44	23	56.1	503	7	CR455290
c 45	23	56.1	519	11	DR19H65

ALIGNMENTS

RESULT 1
LOCUS B2245783 638 bp DNA linear GSS 12-OCT-2002
DEFINITION CH230-325M4.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-325M4, genomic survey sequence.

ACCESSION B2245783
VERSION B2245783.1 GI:2396047
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE
AUTHORS Zhao,S., Shetty,U., Shatsman,S., Tsagaye,G., Geer,K.,
Riggs,F., de Jong,P. and Frazer,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
1 (bases 1 to 638)
Unpublished (1999)

TITLE
JOURNAL
COMMENT
Other GSSs: CH230-325M4.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org

Clones are derived from the rat BAC library CHORI-230
(http://www.choi.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.choi.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
Plate: 325 row: M column: 4
Seq primer: SPE
Clase: BAC ends.

FEATURES

source
1..638
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SNHed/MC9"
/db_xref="taxon:10116"
/clone="CH230-325M4"
/sex="Female"
/cell_type="Brain"

ORIGIN

/clone.lib="CHORI-230 Segment 2"
/note=Vector: PITABAC1.3; Site_1: Mbol; Site_2: Mbol;
CHORI-230 Rat (BN/SENhd/MCW) BAC library produced by
Pieter de Jong"

Query Match 62.9%; Score 25.8; DB 9; Length 638;
Best Local Similarity 93.1%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 9 GTCATATGAGAAAACAAATATTGCGTT 37
Db 287 GTCATATGAGAAAACAAATATTGCGTT 259

RESULT 2
CR750796 583 bp mRNA linear EST 30-AUG-2004
LOCUS DKF2P470M2222.r1 470 (synonym: pliv1) Pongo pygmaeus CDNA clone
DEFINITION DKF2P470M2222.5', mRNA sequence.
ACCESSION CR750796
VERSION CR750796.1 GI:51669963
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pongo.
1 (bases 1 to 583)
Pousatka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Well, B., Amid, C., Osanger, A., Fodo, G., Han, M. and
Wiemann, S.
Pongo pygmaeus mRNA (Pousatka, A., Albert, R., Moosmayer, P., et al.)
Unpublished (2004)
CONTACT: MIPS

TITLE MIPS
JOURNAL Ingolstedter Landstr.1, D-85764 Neuberg, Germany
COMMENT This is the 5' sequence of the clone insert from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email: s.wiemann@dkfz-heidelberg.de; any. Please contact RPD for
ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKF2P470M2222
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source 1..583
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKF2P470M2222"
/issue_type="liver"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="470 (synonym: pliv1)"
/note="Vector: pSPORT1_sfi; Site_1: SfiI; Site_2: SfiIb"

ORIGIN

Query Match 60.0%; Score 24.6; DB 7; Length 563;
Best Local Similarity 76.9%; Pred. No. 3.7e+02;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2 TTAGGAGTCATATGAGAAAACAAATATTGCGTTATA 40
Db 545 TTATGAGGAGATCGATCGAAGAAAATATTGAGTTATA 583

RESULT 3
CR869763 662 bp DNA linear GSS 19-NOV-2004
LOCUS CR869763
DEFINITION Sus scrofa BES, genomic survey sequence.
ACCESSION CR869763
VERSION CR869763.1 GI:55868009
KEYWORDS GSS; Bac-end sequence BES; Genome Survey Sequence.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;
Sus.
1 (bases 1 to 662)
Rogel-Gallard, C., Bourgeaux, N., Billault, A., Vaiman, M. and
Chardon, P.
10449899
Construction of a swine BAC library: application to the
characterization and mapping of porcine type C endoviral elements
Cytogenet. Cell Genet. 85 (3-4), 205-211 (1999)

TITLE JOURNAL
PUBMED 10449899
REFERENCE 2 (bases 1 to 662)
Chardon, P., Iannucci, N., Roig, A., Dossat, C., Demars, J.,
Rogel-Gallard, C., Roy, A., Schibler, L. and Milan, D.
A physical map of the swine genome
Unpublished
3 (bases 1 to 662)
Genoscope.
Direct Submission
Submitted (18-NOV-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

TITLE JOURNAL
AUTHORS Unpublished
REFERENCE 3 (bases 1 to 662)
Genoscope.
Direct Submission
Submitted (18-NOV-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES
source 1..662
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="Large White"
/db_xref="taxon:9823"
/clone="b10733B03"
/sex="male"
/cell_type="fibroblast"
/clone_lib="SBAB"
/note="Genoscope sequence ID : IH0AAA72CG01FM1"

ORIGIN

Query Match 60.0%; Score 24.6; DB 11; Length 662;
Best Local Similarity 76.9%; Pred. No. 3.7e+02;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CTTAGGAGTCATATGAGAAAACAAATATTGCGTTAT 39
Db 111 CTTGGAGAGACAGATCGAAGAAAATCTTAGGTGAT 149

RESULT 4
CV649651 701 bp mRNA linear EST 01-JAN-2005
LOCUS EST928404 Field isolate CDNA library Plasmodium vivax CDNA clone
DEFINITION PVMEF54.5' end, mRNA sequence.
ACCESSION CV649651
VERSION CV649651.1 GI:56956469
KEYWORDS EST.
SOURCE Plasmodium vivax (malaria parasite P. vivax)
ORGANISM Plasmodium vivax
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 701)
Carlton, J.M. and Cui, L.
A survey of genes in Plasmodium vivax by EST sequencing
Unpublished (2004)
CONTACT: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-838-0208
Email: carlton@tigr.org
Seq primer: TL

FEATURES
source 1..701
/organism="Plasmodium vivax"
/mol_type="mRNA"
/strain="field isolate"

/db xref="taxon:5855"
/clone_1ib="Field isolate cDNA library"
/note="Vector: lambda Triplex2; Site_1: Sfi 1A; Site_2: Sfi 1B; Plasmodium vivax field isolate cDNA library made in lambda Triplex2. Inserts cloned unidirectionally in the Sfi 1A and Sfi 1B sites. Mass excision of library produced inserts in Triplex2 plasmid. Inserts sequenced from either 5' or 3' end using Triplex2 sequencing primer or polydr 24 bp primer respectively."

ORIGIN

Query Match 60.0%; Score 24.6; DB 7; Length 701;
Best Local Similarity 87.1%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 AGAGGTCATATGAAAAACAAATATTCGC 34
|||||
DB 379 AGGAGTTCATCTGAAAAACAAAGATTGGC 409
|||||

RESULT 5
CN3635258 771 bp mRNA linear EST 01-JAN-2005
DEFINITION EST914011 Field isolate cDNA library Plasmodium vivax cDNA clone
ACCESSION CV635258
VERSION CV635258
KEYWORDS CV635258.1 GI:56942076
SOURCE EST.
ORGANISM Plasmodium vivax (malaria parasite P. vivax)
Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 771)
Carlton, J.M., and Cui, L.
A survey of genes in Plasmodium vivax by EST sequencing
Unpublished (2004)
Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-838-0208
Email: carlton@tigr.org
Seq primer: Ti.

FEATURES
source
location/Qualifiers
1..771
/organism="Plasmodium vivax"
/mol_type="mRNA"
/strain="Field isolate"
/db_xref="taxon:5855"
/clone="PVMK89"
/clone_1ib="Field isolate cDNA library"
/note="Vector: lambda Triplex2; Site_1: Sfi 1A; Site_2: Sfi 1B; Plasmodium vivax field isolate cDNA library made in lambda Triplex2. Inserts cloned unidirectionally in the Sfi 1A and Sfi 1B sites. Mass excision of library produced inserts in Triplex2 plasmid. Inserts sequenced from either 5' or 3' end using Triplex2 sequencing primer or polydr 24 bp primer respectively."

ORIGIN

Query Match 60.0%; Score 24.6; DB 7; Length 771;
Best Local Similarity 87.1%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 AGAGGTCATATGAAAAACAAATATTCGC 34
|||||
DB 185 AGGAGTTCATCTGAAAAACAAAGATTGGC 215
|||||

RESULT 6
CN361845 500 bp mRNA linear EST 16-MAY-2004
LOCUS CN361845

DEFINITION 17000424189354 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN361845
VERSION CN361845.1 GI:47361779
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 500)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 500 Std Error: 0.00.

FEATURES
source
location/Qualifiers
1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/ligase_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
/clone_1ib="GRN_EB"
/note="Cligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 59.5%; Score 24.4; DB 7; Length 500;
Best Local Similarity 82.4%; Pred. No. 4.3e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGAAAAACAAATATTCGC 35
|||||
DB 326 TTAGAGGTCATATGAAAAACAAATATTCAGCATTCGCTG 359
|||||

RESULT 7
CN361847 521 bp mRNA linear EST 16-MAY-2004
LOCUS CN361847
DEFINITION 17000425249082 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN361847
VERSION CN361847.1 GI:47361781
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 521)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658

FEATURES
source
Fax: 650 473 7760
Email: rbrandenberger@eron.com
Insert Length: 521 Std Error: 0.00.
Location/Qualifiers
1..521

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 7; Length 521;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_id="GRN ES"
/note="Toigo dr primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 7; Length 521;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_id="GRN ES"
/note="Toigo dr primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

RESULT 8
BP350103
LOCUS
DEFINITION
BP350103 Sugano cDNA library, brain Homo sapiens cDNA clone
SZR07565, mRNA sequence.
ACCESSION
BP350103
VERSION
BP350103.1 GI:52280089
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, U., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL
PUBMED
15342556
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SZR07565"
/issue_type="brain"
/clone_id="Sugano cDNA library, brain"

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 3; Length 582;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SZR07565"
/issue_type="brain"
/clone_id="Sugano cDNA library, brain"

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 3; Length 582;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SZR07565"
/issue_type="brain"
/clone_id="Sugano cDNA library, brain"

RESULT 9
AL703924
LOCUS
DEFINITION
AL703924 669 bp mRNA linear EST 04-SEP-2003
DKFZp686D1628_r1 666 (synonym: hlc3) Homo sapiens cDNA clone

ACCESSION
AL703924
VERSION
AL703924.1 GI:19687279
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 669)
Ostenwalder, B., Obermaier, B., Mewes, W., Mewes, H.W., Well, B. and Wiemann, S.
EST (Ostenwalder, B., Obermaier, B., Mewes, H.W., Well, B. and Wiemann, S.)
Unpublished (2001)
JOURNAL
CONTACT: MIPS

FEATURES
source
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D1628"
/dev stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hlc3)"
/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..669

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 1; Length 669;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D1628"
/dev stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hlc3)"
/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..669

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 1; Length 669;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D1628"
/dev stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hlc3)"
/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..669

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 1; Length 669;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D1628"
/dev stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hlc3)"
/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..669

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 1; Length 669;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D1628"
/dev stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hlc3)"
/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..669

ORIGIN
Query Match
Best Local Similarity 82.4%; Score 24.4; DB 1; Length 669;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D1628"
/dev stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hlc3)"
/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..669

6 Taft Court, Suite 100, Rockville, MD 20850, USA
 Tel: 301 340 3188
 Fax: 301 340 8606
 Email: CDNA@origene.com
 This EST submission is part of an on-going human full-length
 cloning project at Origene Technologies, Inc.
 Please contact Origene for access.
 Origene Technologies, Inc.
 6 Taft Ct. Suite 100
 Rockville, MD 20850
 Tel: (301) 340-3188
<http://www.origene.com>
 Seq primer: PCMV6 5prime forward vector primer, Origene
 Technologies Inc.

FEATURES

source

Location/Qualifiers

1..790

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TC114916"

/tissue_type="Fetal Brain"

/clone_lib="Human fetal brain, large insert, PCMV

expression library"

/note="Organ: Fetal Brain; Vector: PCMV6-XL4; Site 1:

ECORI; Site 2: XhoI/Sall compatible end ligatio; Oligo-dT

*, primed reverse transcription optimized for large and GC

rich mRNA transcripts, cDNA size selection, optimized

ligation for large inserts into mammalian expression

vector, random clones selected for end sequence

verification of full-length genes"

ORIGIN

Query Match

Best Local Similarity 59.5%; Score 24.4; DB 8; Length 790;

Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAACAAATATATGCG 35

Db 585 TTAGAGGTCATATGGAACAAATCAAGCATTCG 618

RESULT 11

LOCUS

602713774P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853730 5',

DEFINITION

821 bp mRNA linear EST 15-MAY-2001

ACCESSION

BG756622 mRNA sequence.

VERSION

BG756622.1 GI:14067275

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

NIH-MGC <http://mgs.nci.nih.gov/>.

1 (bases 1 to 821)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM1700 row: d column: 19

High quality sequence stop: 790.

Location/Qualifiers

1..821

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="IMAGE:4853730"
 /tissue_type="Primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"

/note="Organ: B-cells; Vector: pOTR7; Site 1: XhoI;
 Site 2: EORI; cDNA made by oligo-dT priming.
 Directionally cloned into EORI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 59.5%; Score 24.4; DB 2; Length 821;

Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAACAAATATATGCG 35

Db 650 TTAGAGGTCATATGGAACAAATCAAGCATTCG 683

RESULT 12

LOCUS

BO424531 AGENCOURT_7907296 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154690

DEFINITION

5', mRNA sequence.

ACCESSION

BO424531 GI:21119846

VERSION

EST.

KEYWORDS

Homo sapiens

SOURCE

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

NIH-MGC <http://mgs.nci.nih.gov/>.

1 (bases 1 to 905)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM13496 row: c column: 11

High quality sequence stop: 683.

Location/Qualifiers

1..905

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6154690"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_67"

/note="Organ: eye; Vector: PCMV-SPORT6; Site 1: NotI;

Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Library constructed by Life

Technologies."

ORIGIN

Query Match

Best Local Similarity 59.5%; Score 24.4; DB 5; Length 905;

Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAACAAATATATGCG 35

Db 217 TTAGGAGGTGCTATGGAATAATCAAGACATTGCTG 250

RESULT 13
LOCUS BM557407 1113 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT 6561799 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5547856
5', mRNA sequence.
ACCESSION BM557407
VERSION BM557407.1 GI:18799373
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 1113)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgrabbs-remail.nih.gov
Tissue Procurement: ATCC/DCMP/DMP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LLM1225 row: n column: 17
High quality sequence stop: 579.
Location/Qualifiers
1. 1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5547856"
/issue_type="melanotic melanoma"
/lab_host="PH10B (phage-resistant)"
/clone_id="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
.\$
Query Match 59.5%; Score 24.4; DB 3; Length 1113;
Best Local Similarity 82.4%; Pred. No. 4.5e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGGAGGTGCTATGGAATAATCAAGACATTGCTG 35
|||||
576 TTAGGAGGTGCTATGGAATAATCAAGACATTGCTG 609
|||||

RESULT 14
LOCUS AY399307 4191 bp DNA linear GSS 12-DEC-2003
DEFINITION Pan troglodytes HOMO180 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399307
VERSION AY399307.1 GI:39755295
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Pan.
1 (bases 1 to 4191)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,

TITLE Adams,M.D. and Cargill,M.
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4191)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. 4191
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/locus_tag="HOMO180"

ORIGIN
gene

Query Match 59.5%; Score 24.4; DB 10; Length 4191;
Best Local Similarity 82.4%; Pred. No. 3e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGGAGGTGCTATGGAATAATCAAGACATTGCTG 35
|||||
451 TTAGGAGGTGCTATGGAATAATCAAGACATTGCTG 484
|||||

RESULT 15
LOCUS AY399306 4503 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HOMO180 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399306
VERSION AY399306.1 GI:39755295
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 4503)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Interfering nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4503)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. 4503
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/locus_tag="HOMO180"

ORIGIN
gene

Query Match 59.5%; Score 24.4; DB 10; Length 4503;
 Best Local Similarity 82.4%; Pred. No. 5e+02; 6; Indels 0; Gaps 0;
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTGATATGAAAAAATAATATTGCGG 35
 DB 451 TTAGAGGTGCTATGAAAAATCAAGACATTGCTG 484

Search completed: April 7, 2006, 18:17:11
 Job time : 325.989 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:17:47 ; Search time 14.6685 Seconds
(without alignments)
4968.475 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 cttagagagtcatacgtgaagaa.....acaaatattgcgttacc 41

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1/COMB_seq:*
- 2: /cgn2_6/prodata/1/ina/5/COMB_seq:*
- 3: /cgn2_6/prodata/1/ina/6A/COMB_seq:*
- 4: /cgn2_6/prodata/1/ina/6B/COMB_seq:*
- 5: /cgn2_6/prodata/1/ina/H/COMB_seq:*
- 6: /cgn2_6/prodata/1/ina/PCTUS/COMB_seq:*
- 7: /cgn2_6/prodata/1/ina/PP/COMB_seq:*
- 8: /cgn2_6/prodata/1/ina/RE/COMB_seq:*
- 9: /cgn2_6/prodata/1/ina/Backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25.4	62.0	41	US-09-495-406-32	Sequence 32, Appl
2	25.4	62.0	41	US-09-816-028A-46	Sequence 46, Appl
3	25.4	62.0	41	US-10-303-162-46	Sequence 46, Appl
4	25.4	62.0	41	US-10-303-134-46	Sequence 46, Appl
5	25.4	62.0	41	US-10-303-118-46	Sequence 46, Appl
6	25.4	62.0	41	US-10-303-128-46	Sequence 46, Appl
7	24.6	60.0	41	US-09-949-016-12707	Sequence 12707, A
8	24.6	60.0	41	US-09-949-016-17026	Sequence 17026, A
9	23	56.1	444	US-10-012-819-93	Sequence 93, Appl
10	22.4	54.6	451924	US-09-949-016-12896	Sequence 12896, A
11	22.4	54.6	451925	US-09-949-016-17305	Sequence 17305, A
12	22	53.7	601	US-09-949-016-196368	Sequence 196368
13	22	53.7	1401	US-09-328-352-4116	Sequence 4116, Ap
14	22	53.7	1422	US-09-107-532A-2287	Sequence 2287, Ap
15	21.8	53.2	601	US-09-949-016-47720	Sequence 47720, A
16	21.8	53.2	6373	US-08-462-728-1	Sequence 1, Appl
17	21.8	53.2	6373	US-08-461-917-1	Sequence 1, Appl
18	21.8	53.2	6373	US-08-464-436-1	Sequence 1, Appl
19	21.8	53.2	6373	US-08-464-436-1	Sequence 1, Appl
20	21.8	53.2	6375	US-08-168-917-5	Sequence 5, Appl
21	21.8	53.2	6375	US-08-460-510-5	Sequence 5, Appl
22	21.8	53.2	6375	US-08-460-490-5	Sequence 5, Appl
23	21.8	53.2	6375	PCT-US92-00730-5	Sequence 5, Appl
24	21.8	53.2	6375	PCT-US92-00862-5	Sequence 5, Appl

25	21.8	53.2	58829	3	US-09-949-016-13146	Sequence 13146, A
26	21.8	53.2	331814	3	US-09-949-016-12008	Sequence 12008, A
27	21.8	53.2	331814	3	US-09-949-016-17056	Sequence 17056, A
28	21.6	52.7	65848	3	US-09-949-016-12285	Sequence 12285, A
29	21.4	52.2	2838	3	US-09-543-681A-28	Sequence 28, Appl
30	21.4	52.2	53558	3	US-09-949-016-16616	Sequence 16616, A
31	21.4	52.2	63187	3	US-09-949-016-12682	Sequence 12682, A
32	21.4	52.2	63187	3	US-09-949-016-16288	Sequence 16288, A
33	21.4	52.2	197331	3	US-09-949-016-13675	Sequence 13675, A
34	21.4	52.2	197332	3	US-09-949-016-17170	Sequence 17170, A
35	21	51.2	789	6	PCT-US96-05350A-1275	Sequence 1275, Ap
36	21	51.2	837	3	US-09-134-000C-2984	Sequence 2984, Ap
37	21	51.2	1218	3	US-09-543-681A-2852	Sequence 2852, Ap
38	21	51.2	2585	3	US-09-573-080A-8	Sequence 8, Appl
39	21	51.2	26076	3	US-09-949-016-12004	Sequence 12004, A
40	21	51.2	26076	3	US-09-949-016-13041	Sequence 13041, A
41	21	51.2	52496	3	US-09-949-016-16118	Sequence 16118, A
42	21	51.2	52496	3	US-09-949-016-16119	Sequence 16119, A
43	21	51.2	640681	3	US-09-790-988-1	Sequence 1, Appl
44	21	51.2	786431	3	US-09-751-389-3	Sequence 3, Appl
45	21	51.2	1830121	3	US-09-557-884-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-495-406-32
; Sequence 32, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; TITLE OF INVENTION: Ganglioside and Ganglioside Mimics
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
US-09-495-406-32
Query Match
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CY 1 CTTAGAGGTCATATGAAAAAACAATATTCGGG 35
DB 1 CTTAGAGGTCATATGAAAAAAGTTATTCGCTG 35
RESULT 2
US-09-816-028A-46
; Sequence 46, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Ganglioside and Ganglioside Mimics
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/816,028A

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-09-816-028A-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35

RESULT 3
US-10-303-162-46
; Sequence 46, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-162-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35

RESULT 4
US-10-303-134-46
; Sequence 46, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
```

```

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-134-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35

RESULT 5
US-10-303-118-46
; Sequence 46, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-118-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35

RESULT 6
US-10-303-128-46
; Sequence 46, Application US/10303128
; Patent No. 6911337
```



```

; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-00011105
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-128-46

```

```

Query Match          62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 CTTAGAGGTCATATGAAAAACAAATATTTGCGG 35
Db 1 CTTAGAGGTCATATGAAAAAGTATATTTGCTG 35

```

```

RESULT 7
US-09-949-016-12707
; Sequence 12707, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12707
; LENGTH: 190078
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(190078)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12707

```

```

Query Match          60.0%; Score 24.6; DB 3; Length 190078;
Best Local Similarity 76.9%; Pred. No. 12;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 2 TTAGAGGTCATATGAAAAACAAATATTTGCGTTATA 40
Db 45703 TTAGAGGTCATATGATGATTAATTCACAGTTGCGTTATA 45741

```

```

RESULT 8
US-09-949-016-17026
; Sequence 17026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17026
; LENGTH: 190078
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(190078)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17026

```

```

Query Match          60.0%; Score 24.6; DB 3; Length 190078;
Best Local Similarity 76.9%; Pred. No. 12;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 2 TTAGAGGTCATATGAAAAACAAATATTTGCGTTATA 40
Db 45703 TTAGAGGTCATATGATGATTAATTCACAGTTGCGTTATA 45741

```

```

RESULT 9
US-10-012-819-93/C
; Sequence 93, Application US/10012819
; Patent No. 6916615
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; TITLE OF INVENTION: Rain, Jean-Christophe
; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(444)
US-10-012-819-93

```

```

Query Match          56.1%; Score 23; DB 3; Length 444;
Best Local Similarity 74.4%; Pred. No. 25;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

Qy 1 CTTAGAGGTCATATGAAAAACAAATATTTGCGTTAT 39
Db 148 CATAGGTGCTTATGATTAATTAATTCATGCGTTAT 110

```

```
RESULT 10
US-09-949-016-12896
; Sequence 12896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12896
; LENGTH: 451924
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12896
```

```
Query Match          54.6%; Score 22.4; DB 3; Length 451924;
Best Local Similarity 81.2%; Pred. No. 82;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 9 GTCATATGGAAAAACAATAATTCGGGTTATA 40
DB 303773 GTCATATGGAAAAACAATAATTCGGGTTATA 303804
```

```
RESULT 11
US-09-949-016-17305
; Sequence 17305, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17305
; LENGTH: 451925
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17305
```

```
Query Match          54.6%; Score 22.4; DB 3; Length 451925;
Best Local Similarity 81.2%; Pred. No. 82;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 9 GTCATATGGAAAAACAATAATTCGGGTTATA 40
DB 303773 GTCATATGGAAAAACAATAATTCGGGTTATA 303804
```

```
RESULT 12
US-09-949-016-196368
; Sequence 196368, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196368
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196368
```

```
Query Match          53.7%; Score 22; DB 3; Length 601;
Best Local Similarity 73.7%; Pred. No. 60;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 2 TTAGAGGTCATATGAAAAACAATAATTCGGGTTAT 39
DB 517 TTAGGAATCAGTTGGAAAAACAATATTTGCCAATTTGT 554
```

```
RESULT 13
US-09-328-352-4116
; Sequence 4116, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4116
; LENGTH: 1401
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4116
```

```
Query Match          53.7%; Score 22; DB 3; Length 1401;
Best Local Similarity 73.7%; Pred. No. 66;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 2 TTAGAGGTCATATGAAAAACAATAATTCGGGTTAT 39
DB 667 TTAGTTCGTCATATCGAAAAACAATAATGCTGATGT 704
```

```
RESULT 14
US-09-107-532A-2287
; Sequence 2287, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSES: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 2287:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1422 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1..1422
 SEQUENCE DESCRIPTION: SEQ ID NO: 2287:
 US-09-107-532A-2287

Query Match 53.7%; Score 22; DB 3; Length 1422;
 Best Local Similarity 83.3%; Pred. No. 66;
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 AGGTCATATGAAAAACAAATATTGCGGT 36
 DB 378 AGAGCATCTGAAAAACAAATATTGCGTT 407

RESULT 15
 US-09-949-016-47720
 ; Sequence 47720, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47720
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-47720

Query Match 53.2%; Score 21.8; DB 3; Length 601;
 Best Local Similarity 70.7%; Pred. No. 71;
 Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAAATATTGCGGTATAC 41
 DB 28 CTTAGAAAGTCATATGAAAAACAAATATTGCGGTATAC 68

Search completed: April 7, 2006, 18:22:43
 Job time : 18.6685 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:02:11 ; Search time 67.1969 Seconds
(without alignments)
5045.535 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 cttagagagtcatacgaa.....acaataatcgcttatac 41

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA_Main:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	41	3	US-09-211-691-3
2	41	100.0	41	6	US-10-317-773-3
3	41	100.0	41	6	US-10-317-773-3
4	29	70.7	65632	9	US-10-915-740A-49
5	29	70.7	2242716	9	US-10-915-740A-1068
6	25.4	62.0	41	3	US-09-816-028A-46
7	25.4	62.0	41	6	US-10-303-161-46
8	25.4	62.0	41	6	US-10-303-118-46
9	25.4	62.0	41	6	US-10-303-128-46
10	25.4	62.0	41	6	US-10-303-134-46
11	25.4	62.0	41	6	US-10-303-162-46
12	25.4	62.0	41	8	US-10-735-419-46
13	25.4	62.0	41	8	US-10-820-536-46
14	25.4	62.0	41	8	US-10-845-408-46
15	25.4	62.0	41	8	US-10-845-412-46
16	25.4	62.0	41	8	US-10-846-219-46
17	25.4	62.0	41	8	US-10-821-604-46
18	25.4	62.0	41	8	US-10-847-983-46
19	25.4	62.0	41	8	US-10-821-573-46
20	25.4	62.0	41	8	US-10-850-807-46
21	25.4	62.0	41	8	US-10-850-125-46
22	25.4	62.0	41	8	US-10-830-825-46
23	25.4	62.0	41	9	US-10-962-334-46

24	25.4	62.0	41	9	US-10-830-997-46	Sequence 46, Appl
25	25.4	62.0	41	9	US-10-962-235-46	Sequence 46, Appl
26	25.4	62.0	41	9	US-10-961-882-46	Sequence 46, Appl
27	24.4	59.5	7183	5	US-10-198-846-11576	Sequence 11576, A
28	23.2	56.6	529	5	US-09-925-065A-733394	Sequence 733394, A
29	23	56.1	444	5	US-10-012-819-93	Sequence 93, Appl
30	23	56.1	1680	7	US-10-282-122A-22490	Sequence 22490, A
31	23	56.1	1780	3	US-09-895-912A-103	Sequence 103, App
32	23	56.1	4391	9	US-10-929-754-4	Sequence 4, Appl
33	23	56.1	3673778	6	US-10-312-841-2	Sequence 2, Appl
34	22.8	55.6	1095	4	US-09-925-065A-723641	Sequence 723641, Sequence 723642, Sequence 723643, Sequence 3367, Ap
35	22.8	55.6	1095	4	US-09-925-065A-723642	Sequence 3, Appl
36	22.8	55.6	1095	4	US-09-925-065A-723643	Sequence 272562, Sequence 272563, Sequence 272564, Sequence 879344, Sequence 907561, Sequence 876021,
37	22.6	55.1	1538	7	US-10-398-221-3367	
38	22.6	55.1	14012	9	US-09-819-994-3	
39	22.6	55.1	14012	9	US-10-473-340-3	
40	22.4	54.6	528	4	US-09-925-065A-722562	
41	22.4	54.6	528	4	US-09-925-065A-722563	
42	22.4	54.6	528	4	US-09-925-065A-722564	
43	22.4	54.6	528	4	US-09-925-065A-879344	
44	22.4	54.6	528	4	US-09-925-065A-907561	
45	22.4	54.6	578	4	US-09-925-065A-876021	

ALIGNMENTS

RESULT 1
US-09-211-691-3
Sequence 3, Application US/09211691
Patent No. US20020034805A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 019957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:SYNTH-P1 5'
US-09-211-691-3
Query Match
Best Local Similarity 100.0%; Score 41; DB 3; Length 41;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTAGAGGTCATATGAAAAAACAATATATTCGGTTATAC 41
DB 1 CTTAGAGGTCATATGAAAAAACAATATATTCGGTTATAC 41
RESULT 2
US-10-317-773-3
Sequence 3, Application US/10317773
Publication No. US20030180928A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
APPLICANT: National Research Council of Canada
TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNAc 4' Epimerase and a

*No Jhl
Cloning Patenting
different fusion
protein*

TITLE OF INVENTION: GalNAc Transferase
FILE REFERENCE: 019633-000812US
CURRENT APPLICATION NUMBER: US/10/317,773
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence:SYNTH-Fl 5'
OTHER INFORMATION: primer
US-10-317-773-3

Query Match 100.0%; Score 41; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAGAAACAAATATTCGGTTATAC 41
DB 1 CTTAGAGGTCATATGAGAAACAAATATTCGGTTATAC 41

RESULT 3
US-10-317-428-3
Sequence 3, Application US/10317428
Publication No. US20030186414A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
APPLICANT: National Research Council of Canada
TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
FILE REFERENCE: 019633-000811US
CURRENT APPLICATION NUMBER: US/10/317,428
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence:SYNTH-Fl 5'
OTHER INFORMATION: primer
US-10-317-428-3

same as above

Query Match 100.0%; Score 41; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAGAAACAAATATTCGGTTATAC 41
DB 1 CTTAGAGGTCATATGAGAAACAAATATTCGGTTATAC 41

RESULT 4
US-10-915-740A-49
Sequence 49, Application US/10915740A
Publication No. US20050191316A1
GENERAL INFORMATION:
APPLICANT: Frazer, Claire M.
APPLICANT: Hickey, Erin
APPLICANT: Peterson, Jeremy

APPLICANT: Tetteijn, Herve
APPLICANT: Venter, J. Craig
APPLICANT: Maignani, Vega
APPLICANT: Galeotti, Cesira
APPLICANT: Mora, Manroa
APPLICANT: Ratti, Giulio
APPLICANT: Scarselli, Maria
APPLICANT: Scarlato, Vincenzo
APPLICANT: Rappuoli, Rino
APPLICANT: Pizze, Mariagrazia
APPLICANT: Grandi, Guido
TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
FILE REFERENCE: 002441.00090
CURRENT APPLICATION NUMBER: US/10/915,740A
CURRENT FILING DATE: 2004-08-11
PRIOR APPLICATION NUMBER: 09/806,866
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: USSN 60/103,794
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: USSN 60/132,068
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: PCT/US99/25373
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 1068
SOFTWARE: PatentIn version 3.2
SEQ ID NO 49
LENGTH: 65632
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-915-740A-49

Query Match 70.7%; Score 29; DB 9; Length 65632;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAGAAACAAATATTCGGTTATAC 41
DB 55054 TATGAGAAACAAATATTCGGTTATAC 55082

RESULT 5
US-10-915-740A-1068/c
Sequence 1068, Application US/10915740A
Publication No. US20050191316A1
GENERAL INFORMATION:
APPLICANT: Frazer, Claire M.
APPLICANT: Hickey, Erin
APPLICANT: Peterson, Jeremy
APPLICANT: Tetteijn, Herve
APPLICANT: Venter, J. Craig
APPLICANT: Maignani, Vega
APPLICANT: Galeotti, Cesira
APPLICANT: Mora, Manroa
APPLICANT: Ratti, Giulio
APPLICANT: Scarselli, Maria
APPLICANT: Scarlato, Vincenzo
APPLICANT: Rappuoli, Rino
APPLICANT: Pizze, Mariagrazia
APPLICANT: Grandi, Guido
TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
FILE REFERENCE: 002441.00090
CURRENT APPLICATION NUMBER: US/10/915,740A
CURRENT FILING DATE: 2004-08-11
PRIOR APPLICATION NUMBER: 09/806,866
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: USSN 60/103,794
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: USSN 60/132,068
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: PCT/US99/25373
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 1068
SOFTWARE: PatentIn version 3.2

SEQ ID NO 1068
LENGTH: 2242716
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-915-740A-1068g

Query Match 70.7%; Score 29; DB 9; Length 2242716;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAACAATATTGCGGTATAC 41
DB 79104 TATGAAAAACAATATTGCGGTATAC 79076

RESULT 6
US-09-816-028A-46
Sequence 46, Application US/09816028A
Patent No. US20020042369A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/09/816,028A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-09-816-028A-46

Query Match 62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35

RESULT 7
US-10-303-161-46
Sequence 46, Application US/10303161
Publication No. US20030148459A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/303,161
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-161-46

Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35

RESULT 8
US-10-303-118-46
Sequence 46, Application US/10303118
Publication No. US20030157655A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/303,118
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-118-46

Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35

RESULT 9
US-10-303-128-46
Sequence 46, Application US/10303128
Publication No. US20030157656A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/303,128
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213

```
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-128-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 10
US-10-303-134-46
; Sequence 46, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-134-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 11
US-10-303-162-46
; Sequence 46, Application US/10303162
; Publication No. US20030157658A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
```

```
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-162-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 12
US-10-735-419-46
; Sequence 46, Application US/10735419
; Publication No. US20040180406A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-735-419-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 13
US-10-820-536-46
; Sequence 46, Application US/10820536
; Publication No. US20040203103A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
```


APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/820,536
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: 10/303,128
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
US-10-820-536-46

Query Match
Best Local Similarity 62.0%; Score 25.4; DB 8; Length 41;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAAGTTATTATGCTG 35

RESULT 14
US-10-845-408-46
Sequence 46, Application US/10845408
Publication No. US20040203112A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/845,408
CURRENT FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
US-10-845-408-46

Query Match
Best Local Similarity 62.0%; Score 25.4; DB 8; Length 41;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAAGTTATTATGCTG 35

RESULT 15
US-10-845-412-46
Sequence 46, Application US/10845412
Publication No. US20040203113A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/845,412
CURRENT FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/10/303,128
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
US-10-845-412-46

Query Match
Best Local Similarity 62.0%; Score 25.4; DB 8; Length 41;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAAGTTATTATGCTG 35

Search completed: April 7, 2006, 18:56:37
Job time : 70.1969 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:05:06 ; Search time 47.8632 Seconds
(without alignments)
3426.906 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41
Sequence: 1 cttgaggagcattcgtgaaa.....acaaatattcggtttac 41

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*

- 1: /SID55/ptodata/1/pubpna/US08_NEW_PUB.seq:**
- 2: /SID55/ptodata/1/pubpna/US06_NEW_PUB.seq:**
- 3: /SID55/ptodata/1/pubpna/US07_NEW_PUB.seq:**
- 4: /SID55/ptodata/1/pubpna/PCF_NEW_PUB.seq:**
- 5: /SID55/ptodata/1/pubpna/US09_NEW_PUB.seq:**
- 6: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq1:**
- 7: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq1:**
- 8: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq2:**
- 9: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq3:**
- 10: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq:**
- 11: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq2:**
- 12: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq3:**
- 13: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq4:**
- 14: /SID55/ptodata/1/pubpna/US60_NEW_PUB.seq:**
- 15: /SID55/ptodata/1/pubpna/US60_NEW_PUB.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.2	56.6	529	US-09-925-065A-733394	Sequence 733394,
2	22.8	55.6	1095	US-09-925-065A-723641	Sequence 723641,
3	22.8	55.6	1095	US-09-925-065A-723642	Sequence 723642,
4	22.8	55.6	1095	US-09-925-065A-723643	Sequence 723643,
5	22.6	55.1	187745	US-11-121-086-83	Sequence 83, Appl
6	22.6	54.6	528	US-09-925-065A-272562	Sequence 272562,
7	22.4	54.6	528	US-09-925-065A-272563	Sequence 272563,
8	22.4	54.6	528	US-09-925-065A-272564	Sequence 272564,
9	22.4	54.6	533	US-10-301-480-349856	Sequence 349856,
10	22.4	54.6	533	US-10-301-480-349857	Sequence 349857,
11	22.4	54.6	533	US-10-301-480-349858	Sequence 349858,
12	22.4	54.6	533	US-10-301-480-349859	Sequence 349859,
13	22.4	54.6	533	US-10-301-480-349860	Sequence 349860,
14	22.4	54.6	533	US-10-301-480-349861	Sequence 349861,
15	22.4	54.6	533	US-10-301-480-349862	Sequence 349862,
16	22.4	54.6	533	US-10-301-480-349863	Sequence 349863,
17	22.4	54.6	533	US-10-301-480-349864	Sequence 349864,
18	22.4	54.6	533	US-10-301-480-349865	Sequence 349865,
19	22.4	54.6	533	US-10-301-480-349866	Sequence 349866,
20	22.4	54.6	533	US-10-301-480-349867	Sequence 349867,
21	22.4	54.6	533	US-10-301-480-349868	Sequence 349868,
22	22.4	54.6	533	US-10-301-480-349869	Sequence 349869,
23	22.4	54.6	533	US-10-301-480-349870	Sequence 349870,
24	22.4	54.6	533	US-10-301-480-349871	Sequence 349871,
25	22.4	54.6	533	US-10-301-480-349872	Sequence 349872,
26	22.4	54.6	533	US-10-301-480-349873	Sequence 349873,
27	22.4	54.6	533	US-10-301-480-349874	Sequence 349874,
28	22.4	54.6	533	US-10-301-480-349875	Sequence 349875,
29	22.4	54.6	533	US-10-301-480-349876	Sequence 349876,
30	22.4	54.6	533	US-10-301-480-349877	Sequence 349877,
31	22.4	54.6	533	US-10-301-480-349878	Sequence 349878,
32	22.4	54.6	533	US-10-301-480-349879	Sequence 349879,
33	22.4	54.6	533	US-10-301-480-349880	Sequence 349880,
34	22.4	54.6	533	US-10-301-480-349881	Sequence 349881,
35	22.4	54.6	533	US-10-301-480-349882	Sequence 349882,
36	22.4	54.6	533	US-10-301-480-349883	Sequence 349883,
37	22.4	54.6	533	US-10-301-480-349884	Sequence 349884,
38	22.4	54.6	533	US-10-301-480-349885	Sequence 349885,
39	22.4	54.6	533	US-10-301-480-349886	Sequence 349886,
40	22.4	54.6	533	US-10-301-480-349887	Sequence 349887,
41	22.4	54.6	533	US-10-301-480-349888	Sequence 349888,
42	22.4	54.6	533	US-10-301-480-349889	Sequence 349889,
43	22.4	54.6	533	US-10-301-480-349890	Sequence 349890,
44	22.4	54.6	533	US-10-301-480-349891	Sequence 349891,
45	22.4	54.6	533	US-10-301-480-349892	Sequence 349892,

ALIGNMENTS

```
RESULT 1
US-09-925-065A-733394/c
; Sequence 733394, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 733394
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-733394

Query Match      56.6% Score 23.2; DB 6; Length 529;
Beech Local Similarity 77.8%; Pred. No. 72;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      4 AGGAGTCATATGCAAAAATATTCGCTTAT 39
Db      319 AGGCATCATATGCAAAAATTTATATGAGTAT 284

RESULT 2
US-09-925-065A-723641/c
; Sequence 723641, Application US/09925065A
```

Publication No. US20040181048A1

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIORITY FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 723641
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723641
```

Query Match 55.6%; Score 22.8; DB 6; Length 1095;
Best Local Similarity 92.3%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGGTCATATGAAAAACAAATATTG 32
DB 435 ACGGCATATGAAAAACAAATATTG 410

RESULT 3

```
; Sequence 723642, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIORITY FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 723642
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723642
```

Query Match 55.6%; Score 22.8; DB 6; Length 1095;
Best Local Similarity 92.3%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGGTCATATGAAAAACAAATATTG 32
DB 435 ACGGCATATGAAAAACAAATATTG 410

RESULT 4

```
US-09-925-065A-723643/c
; Sequence 723643, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIORITY FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 723643
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723643
```

Query Match 55.6%; Score 22.8; DB 6; Length 1095;
Best Local Similarity 92.3%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGGTCATATGAAAAACAAATATTG 32
DB 435 ACGGCATATGAAAAACAAATATTG 410

RESULT 5

```
US-11-121-086-83
; Sequence 83, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIORITY FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patencin version 3.3
; SEQ ID NO 83
; LENGTH: 187745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-83
```

Query Match 55.1%; Score 22.6; DB 14; Length 187745;
Best Local Similarity 86.2%; Pred. No. 7.9e+02;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GGTGATATGAAAAACAAATATTGCGGT 36
DB 66722 GGTGATATGAAAAACAAATATTGCGT 66750

RESULT 6

```
US-09-925-065A-272562
; Sequence 272562, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-24
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-272562

Query Match          54.6%; Score 22.4; DB 6; Length 528;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

QY 1 CTTAGAGGTCATATGAGAAAACAAATATTCGGCTTATA 40
DB 322 CTTAGAGATCATATCTCAATCAATCAATATTCAGCCTTAGA 361
```

```

RESULT 7
US-09-925-065A-272563
; Sequence 272563, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-24
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-272563
```

```

Query Match          54.6%; Score 22.4; DB 6; Length 528;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

QY 1 CTTAGAGGTCATATGAGAAAACAAATATTCGGCTTATA 40
DB 322 CTTAGAGATCATATCTCAATCAATCAATATTCAGCCTTAGA 361

RESULT 8
US-09-925-065A-272564
```

```

; Sequence 272564, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-24
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-272564

Query Match          54.6%; Score 22.4; DB 6; Length 528;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

QY 1 CTTAGAGGTCATATGAGAAAACAAATATTCGGCTTATA 40
DB 322 CTTAGAGATCATATCTCAATCAATCAATATTCAGCCTTAGA 361
```

```

RESULT 9
US-10-301-480-349856
; Sequence 349856, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-349856
```

```

Query Match          54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

QY 1 CTTAGAGGTCATATGAGAAAACAAATATTCGGCTTATA 40
DB 327 CTTAGAGATCATATCTCAATCAATCAATATTCAGCCTTAGA 366

RESULT 10
US-10-301-480-349857
; Sequence 349857, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 349857
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-349857

Query Match          54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGCAAAACAAATATTCGCGTTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 11
US-10-301-480-349858
; Sequence 349858, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 349858
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-349858

Query Match          54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGCAAAACAAATATTCGCGTTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 12
US-10-301-480-963265
; Sequence 963265, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
```

```

; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 963265
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-963265

Query Match          54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGCAAAACAAATATTCGCGTTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 13
US-10-301-480-963266
; Sequence 963266, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 963266
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-963266

Query Match          54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGCAAAACAAATATTCGCGTTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 14
US-10-301-480-963267
; Sequence 963267, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 963267
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-963267
```

Query Match 54.6%; Score 22.4; DB 10; Length 533;
 Best Local Similarity 72.5%; Pred. No. 1.4e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAAATATTGCGTTATA 40
 DB 327 CTTAGAGGTCATATCTCAATCAATCAATATTAGCCTTAGA 366

RESULT 15

US-09-925-065A-879344
 / Sequence 879344, Application US/09925065A
 / Publication No. US20040181048A1
 / GENERAL INFORMATION:
 / APPLICANT: Wang, David G.
 / TITLE OF INVENTION: Identification and Mapping of Single
 / TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
 / FILE REFERENCE: 108827.135
 / CURRENT APPLICATION NUMBER: US/09/925,065A
 / PRIOR FILING DATE: 2001-08-08
 / PRIOR APPLICATION NUMBER: US 60/243,096
 / PRIOR FILING DATE: 2000-10-24
 / PRIOR APPLICATION NUMBER: US 60/252,147
 / PRIOR FILING DATE: 2000-11-20
 / PRIOR APPLICATION NUMBER: US 60/250,092
 / PRIOR FILING DATE: 2000-11-30
 / PRIOR APPLICATION NUMBER: US 60/261,766
 / PRIOR FILING DATE: 2001-01-16
 / PRIOR APPLICATION NUMBER: US 60/289,846
 / PRIOR FILING DATE: 2001-05-09
 / NUMBER OF SEQ ID NOS: 957086
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 879344
 / LENGTH: 575
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-925-065A-879344

Query Match 54.6%; Score 22.4; DB 6; Length 575;
 Best Local Similarity 72.5%; Pred. No. 1.5e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGAAAAACAAATATTGCGTTATAC 41
 DB 223 TTAGAGGTCATATGAAAAACAAATATTGCGTTATAC 262

Search completed: April 7, 2006, 21:23:54
 Job time : 48.8632 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:52 ; Search time 307.615 Seconds
(without alignments)
8315.447 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45
Sequence: 1 cgacagaaatccgcacacgc.....ctcgtgataagaatgttttc 45

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	27	60.0	687	1	NM060146
C 2	27	60.0	1875	1	U04328
C 3	27	60.0	4176	1	NM85YN
C 4	27	60.0	4908	1	NM95053
C 5	27	60.0	110000	1	AE002098
C 6	27	60.0	349980	6	AX043922
C 7	25.8	57.3	239048	14	AC094874
C 8	25	55.6	2603	8	BC016953
C 9	24.6	54.7	152506	9	AC145589
C 10	24.6	54.7	166529	9	AC147621
C 11	24.2	53.8	185020	14	AC107117
C 12	24.2	53.3	367	6	COS05083
C 13	24	53.3	367	6	COS10708
C 14	24	53.3	237515	14	AC126059
C 15	24	52.9	174372	8	AC099051
C 16	23.8	52.9	217409	8	AC123901
C 17	23.8	52.9	227693	14	AC095313
C 18					

C 19	23.8	52.9	233866	14	AC128261	AC128261 Rattus no
C 20	23.8	52.9	239037	14	AC135695	AC135695 Rattus no
C 21	23.8	52.9	271667	14	AC160572	AC160572 Bos tauru
C 22	23.8	52.9	318143	14	AC095005	AC095005 Rattus no
C 23	23.6	52.4	1174	1	CQ895926	CQ895926 Sequence
C 24	23.6	52.4	1174	1	MC150	Z33105 M.capricolu
C 25	23.6	52.4	257159	14	AC094460	AC094460 Rattus no
C 26	23.6	52.4	306172	14	AC103326	AC103326 Rattus no
C 27	23.4	52.0	351	6	CQ460253	CQ460253 Sequence
C 28	23.4	52.0	943	2	AF508884	AF508884 Xyleboru
C 29	23.4	52.0	2759	15	AY436553	AY436553 Theilingi
C 30	23.4	52.0	3812	6	CQ850365	CQ850365 Sequence
C 31	23.4	52.0	3812	8	AK127490	AK127490 Homo sapi
C 32	23.4	52.0	4470	6	CQ841515	CQ841515 Sequence
C 33	23.4	52.0	4470	8	AK124570	AK124570 Homo sapi
C 34	23.4	52.0	55151	15	AB020751	AB020751 Arabidops
C 35	23.4	52.0	110000	14	AC153053	Continuation (2 of
C 36	23.4	52.0	183827	8	AC005899	AC005899 Homo sapi
C 37	23.4	52.0	184841	14	AC129486	AC129486 Homo sapi
C 38	23.2	51.6	720	6	AR397330	AR397330 Sequence
C 39	23.2	51.6	132948	8	HS349A12	AL033520 Human DNA
C 40	23.2	51.6	148259	9	AC104059	AC104059 Mus muscu
C 41	23.2	51.6	175614	14	AC159570	AC159570 Callithri
C 42	23.2	51.6	181669	9	AL645903	AL645903 Mouse DNA
C 43	23.2	51.6	205647	9	AC111132	AC111132 Mus muscu
C 44	23.2	51.6	228380	14	AC133352	AC133352 Rattus no
C 45	23.2	51.6	244451	14	AC125765	AC125765 Rattus no

ALIGNMENTS

RESULT 1
LOCUS NM060146/c 687 bp DNA linear BCT 11-JUN-1997
DEFINITION Neisseria meningitidis CMP-sialic acid synthetase gene, complete cds.
ACCESSION U60146
VERSION U60146.1 GI:1549340
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
REFERENCE 1 (bases 1 to 687)
AUTHORS Gilbert,M., Watson,D.C. and Makarchuk,W.W.
TITLE Purification and characterization of the recombinant CMP-sialic acid synthetase from Neisseria meningitidis
JOURNAL Biochemol. Lett. 19, 417-420 (1997)
JOURNAL 2 (bases 1 to 687)
AUTHORS Gilbert,M., Watson,D.C. and Makarchuk,W.W.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1996) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada

FEATURES
source
CDS
1..687
/location=Qualifiers
1..687
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="406Y; NRC04030"
/db_xref="taxon:487"
/note="capsule type: Y; LPS type: U3"
1..687
/note="first 40 residues of the N-terminus of the recombinant product were determined experimentally."
/evidence=experimental
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AA060780.1"
/db_xref="GI:1549341"
/translation="MEKONIAVILARQNSKGLPKNLKRNKNGISLGHITINAISKC
FDRIIVSDGGLIAEBKNNFVEVVLRPAELASPTASISGVTHALERTIGNSGVTL

ORIGIN

Query Match 60.0%; Score 27; DB 1; Length 687;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 GCTTCTTGATTAAGATGTTTC 45
|||||
684 GCTTCTTGATTAAGATGTTTC 658

RESULT 2
U04328/c 1875 bp DNA linear BCT 15-APR-1994
DEFINITION Neisseria meningitidis NMB CMP-N-acetylneuraminic acid synthetase (synB) and synX (synX) genes, complete cds.
ACCESSION U04328
VERSION U04328.1 GI:460144
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
REFERENCE
AUTHORS Swartley, J.S. and Stephens, D.S.
TITLE Identification of a genetic locus involved in the biosynthesis of N-acetyl-D-mannosamine, a precursor of the (alpha 2-->8)-linked polysialic acid capsule of serogroup B Neisseria meningitidis
J. Bacteriol. 176 (5), 1530-1534 (1994)
8113198
2 (bases 1 to 1875)
REFERENCE
AUTHORS Stephens, D.S.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-1993) David S. Stephens, Emory University, Department of Medicine, 69 Butler Street, Atlanta, GA 30303, USA
Location/Qualifiers
1. 1875
/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"
/strain="NMB"
/isolate="W7"
/db_xref="taxon:487"
/note="serogroup B"
52. 1185
/gene="synX"
52. 1185
/gene="synX"
/function="proposed to be involved in the biosynthesis of N-acetyl-D-mannosamine"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="SynX"
/protein_id="AAI17654.1"
/db_xref="GI:460145"
/translation="MKRIICITGRADFGKLPKLAAYIENHDPLEHLIVTGMMNMT YGRIVKRENTYOHNTYLFNSOIGEPGAVLGNITTFISRLSDEIEPDVMIHGRLL BALAGAAGALSRLVCHIEGELSGVDSDIRHSISKSHILVANBOAVRIYVOMG EKAKHIIIGSPDLVMAASTLPISLEVEYEGGLPRNNGISMFHYTTEAHIMPOYA AQYFKALEISGQNIISIPNNDGTESIIOELKIOSDKFIAPPSIRFEYFVYLKHA KFWGNSSAGIREAPLYGPSIDVGTROSNRHMGSKIHTDYETKNIIPDAIQOACSLG KFEADITFNGCDRTSTERPAEVIINPWNVSAOKRFIDLN"
1189. 1875
/gene="synB"
1189. 1875
/gene="synB"
/standard_name="CMP-NANA synthetase"
/codon_start=1
/transl_table=11
/product="CMP-N-acetylneuraminic acid synthetase"

ORIGIN

Query Match 60.0%; Score 27; DB 1; Length 1875;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 GCTTCTTGATTAAGATGTTTC 45
|||||
1872 GCTTCTTGATTAAGATGTTTC 1846

RESULT 3
NMBSYN/c
LOCUS NMBSYN
DEFINITION N meningitidis (group B) ctrA, synA, synB, synC and sialyltransferase genes.
ACCESSION X78068
VERSION X78068.1 GI:530037
KEYWORDS
SOURCE CMP-NeuNAC synthetase; ctrA gene; sialyltransferase; synA gene; synB gene; synC gene.
ORGANISM Neisseria meningitidis serogroup B
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
REFERENCE
AUTHORS Ganguli, S., Zapata, G., Wallis, T., Reid, C., Boulnois, G., Vann, W.F. and Roberts, I.S.
TITLE Molecular cloning and analysis of genes for sialic acid synthesis in Neisseria meningitidis group B and purification of the meningococcal CMP-NeuNAC synthetase enzyme
J. Bacteriol. 176 (15), 4583-4589 (1994)
8045888
2 (bases 1 to 4176)
REFERENCE
AUTHORS Ganguli, S.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-1994) S. Ganguli, University of Leicester, Dept of Microbiology, PO Box 138, Medical Sciences Building, University Road, Leicester LE1 9HN, UK
Related sequences: M57677 & M64389.
Location/Qualifiers
1. 4176
/organism="Neisseria meningitidis serogroup B"
/mol_type="genomic DNA"
/isolate="NCTC8249"
/db_xref="taxon:491"
complement(<1. 657)
/gene="ctrA"
complement(<1. 657)
/gene="ctrA"
792. 1925
/gene="synA"
/gene="synA"
792. 1925
/transl_table=11
/codon_start=1
/protein_id="CAA54982.1"
/db_xref="GI:530038"
/db_xref="GOA:Q57141"
/db_xref="UniProt:IPR003331"
/db_xref="TrEMBL:Q57141"
/translation="MKRIICITGRADFGKLPKLAAYIENHDPLEHLIVTGMMNMT YGRIVKRENTYOHNTYLFNSOIGEPGAVLGNITTFISRLSDEIEPDVMIHGRLL BALAGAAGALSRLVCHIEGELSGVDSDIRHSISKSHILVANBOAVRIYVOMG EKAKHIIIGSPDLVMAASTLPISLEVEYEGGLPRNNGISMFHYTTEAHIMPOYA AQYFKALEISGQNIISIPNNDGTESIIOELKIOSDKFIAPPSIRFEYFVYLKHA KFWGNSSAGIREAPLYGPSIDVGTROSNRHMGSKIHTDYETKNIIPDAIQOACSLG

	gene	KFEADPTENGSDTSTSTERFAEVINNPBETWNVSAQKGFIDNLN"
CDS	1929..2615 /gene="synb" 1929..2615 /gene="synb" /transl_table=1 /product="CMP-NeuNac synthetase" /db_xref="GI:530039"	
gene	/db_xref="GOA:P0A027" /db_xref="InterPro:IPR003329" /translacion="MERONITAVILARONSCKLPKCNRYNGISLIGHTINAASSKC PDRIVSTDGGLIABEAKNRGEVYVLPAELASTPASISGVIALHATTISNSGTVL LEQPSPLRTGAHIREAFSLFDPEKIKGSVNSACPMENHP LKTLIQINGEYAAPMHSLS LOPRROOLPQAIFPNCGAIYINDTASLIANNCFPIAPTLYIMSHODSIDDTIELDIQQ AENTLNKHES"	
CDS	2616..3665 /gene="sync" 2616..3665 /gene="sync" /codon_start=1 /transl_table=1 /protein_id="CA54984.1" /db_xref="GI:530040" /db_xref="GOA:Q7DDU0" /db_xref="InterPro:IPR004144" /db_xref="InterPro:IPR006014" /db_xref="InterPro:IPR006190" /db_xref="UniProt/TREMBL:Q7DDU0" /translation="MNUNNEFKIGNRSVGNIHEPLLICEIGINHESGLKTAFEWDAA YNMAFISTPFRAALRLORMDIPAYKIGSGECNNYLILKI VASFGRPIILSTGMSI ESIKSVEYIIIRBAGVPALLHCTNIYTPPYEDVALGGANDLSAEPALIIGLSDHTD NYACLGAVALGGSLIERHFTRDMRPDPDVIYSWNPDTEKEIKOGAHATKLARGCKD TIINGERPTDKPAFASVYADKDIOKGELLSGDNLMYKRFPGNDFSVNETYLLFGKAAA CNIRKAQIKKTIDLE"	
CDS	3749..>4176 /codon_start=1 /transl_table=1 /product="galyttransferase" /protein_id="CA54985.1" /db_xref="GI:530041" /db_xref="GOA:Q79AE7" /db_xref="InterPro:IPR010866" /db_xref="UniProt/TREMBL:Q79AE7" /translation="MLKKIKKALKQPKKFPDDSSMWLTSPFYLLPPRNNTLVISNQG LNUQSILKIKQLTNLNLVILIYISKNLMPKLFVHQSNKUDFBIYIFELPRSPNNT PKKLVIYRSYRKLIINTIQPHALYMLSFTHSYLSIAK" 3773..3790 /note="putative"	
stem_loop		
ORIGIN		
Query Match	60.0%; Score 27; DB 1; Length 4176;	
Best Local Similarity	100.0%; Pred. No. 8.9;	
Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	19 GCTTCCTTGATTAAGAAGTTTTC 45 	
Dn	2612 GCTTCCCTTGATTAAGAAGTTTTT 2586 	
RESULT 4		
NMM95053/c	NMM95053 4908 bp DNA linear BCT 04-APR-2002	
LOCUS	Neisseria meningitidis Sia (sia), CMP-NeuNAc synthetase (siab),	
DEFINITION	siaC (siaC), and siad (siad) genes, complete cds.	
ACCESSION	M95053 M64289 M96562	
VERSION	M95053.1 GI:520732	
KEYWORDS	siaA; siab; siaC; siad. Neisseria meningitidis	
SOURCE	Neisseria meningitidis	
ORGANISM		

REFERENCE	Bacteria: Proteobacteria; Betaproteobacteria; Neisseriales;
AUTHORS	Neisseriaceae; Neisseria.
TITLE	1 (bases 2861 to 4641) Frosch,M., Edwards,U., Bousset,K., Krause,B. and Weisgerber,C. Evidence for a common molecular origin of the capsule gene loci in gram-negative bacteria expressing group II capsular polysaccharides <i>Mol. Microbiol.</i> 5 (5), 1251-1263 (1991)
JOURNAL	1659649
PUBMED	2 (bases 1210 to 1987) Edwards,U. and Frosch,M. Sequence and functional analysis of the cloned <i>Neisseria meningitidis</i> CMP-NeuNac synthetase <i>FEMS Microbiol. Lett.</i> 75 (2-3), 161-166 (1992)
REFERENCE	1398032
AUTHORS	3 (bases 1 to 4908) Edwards,U., Muller,A., Hammerschmidt,S. and Frosch,M. Molecular analysis of the biosynthesis pathway of the alpha 2,8 polysialic acid capsule by <i>Neisseria meningitidis</i> serogroup B unpublished
JOURNAL	4 (bases 1210 to 1987)
REFERENCE	Frosch,M. Direct Submission Submitted (28-OCT-1992) Matthias Frosch, Medizinische Hochschule Hannover, Institut fur Medizinische Mikrobiologie, Konstanty Gutschow Strasse 8, Hannover, Germany, 30623
AUTHORS	5 (bases 1 to 4908) Frosch,M. Direct Submission Submitted (25-MAY-1994) Matthias Frosch, Medizinische Hochschule Hannover, Institut fur Medizinische Mikrobiologie, Konstanty Gutschow Strasse 8, Hannover, Germany, 30623
JOURNAL	On or before Apr 4, 2002 this sequence version replaced gi:150354, gi:497288.
COMMENT	
FEATURES	Location/Qualifiers
SOURCE	1..4908 /organism="Neisseria meningitidis" /mol_type="genomic DNA" /isolate="B1940" /specific_host="Homo sapiens" /db_xref="taxon:487" 51..56 /standard_name="-10 signal"
gene	174..1307 /gene="siaA" 174..1307 /gene="siaA" /codon_start=1 /transl_table=11 /product="SiaA" /protein_id="AA20475.1" /db_xref="GI:520733" /translation="MKRIILCTGTGRADFGKLPDLAYIENPDLHLIVTGAMPKT YGTVAGVRENYOHTYIPSNQICGBMGAVLGNTITETSRSDTEBPWMVTHGRL BALAGAVALSSRLVCHIBEGELSGTVDSIRHSISKSHITLVANEQAVITLVONG EKKRHIIIGSPLDVAASSTPLSLBEVEYGLPYENVGI SNEHPVTTEALMPDYA AQCFKLAEISGCNNIIISIPNDNGTESIIQLKLYOSDKFIAPSIREFEYLTKKA KFMWGSAGIRRAPLYGPSIDVGTRONNRHGKSI IHTDYEYKNI PDAIQACSIG KPFAADTPNGGDRTTSTERAFYINNPEYTNVSAQRFDLNL"
CDS	1311..1997 /gene="siab" 1311..1997 /gene="siab" /ciliation=121 /codon_start=1 /transl_table=11 /product="CMP-NeuNac synthetase" /protein_id="AA20476.1" /db_xref="GI:520734" /translation="MEKONIATVILARNSKGPLEKNIARKNGISILGHTINAISSKC FDIILYSTGGLLIAERAKNGGVAVLRPAIADTSASISGYHALETIGNSGTYTL LGPTSPRTCAHI REAFSLFEDEIKSYVSACMEHRPIKTLLQINGRVARPRHSD LEPDRPOLPAFRPNGALITYINDIASLIANNCFPIATPKLYIMSHODSIDIDELDQQ AENILNHKES"
gene	
CDS	
gene	

```
gene 1998. 3047
      /gene="biac"
CDS 1998. 3047
      /gene="biac"
      /citation=[1]
      /codon_start=1
      /transl_table=1
      /product="biac"
      /protein_id="AAA20477.1"
      /db_xref="GI:520735"
      /translation="MNNNEFKIKGRSVGNHEPLICEIGINHGSLKTAPEWYDA
      VYAGAEVGHQTHIVEDWSDAKQVIFPNADVSLEYEIMERCALNEEDITKXEYVES
      KMIFISPTSPSAALRIQRMIDIPAKIGSGCANNYPILKIVASGKPTILSTGNSTI
      ESIKKSVETIRAGVPYALHCTNITYTPPYEVRGAMNDLSAPPDALIGSDITLD
      NYACGVALGSGSILERTHTDMRDPRGPVIVCSMPDTPKEIKQGHAKLAKGKGKD
      TIIAEKPTKDPAFASVADKDIKKGELLGSLNWKVRGNDFSVNEYETLLGRVAA
      CNIRGAQIKKIDIE"
      3131. 4618
      /gene="biad"
      3131. 4618
      /gene="biad"
      /citation=[3]
      /transl_table=1
      /codon_start=1
      /protein_id="AAA20478.1"
      /db_xref="GI:520736"
      /translation="MKIKTKKALFQPKKFFQDSMWLTSPFYLTPPRNNLPYISNLQ
      LKAVOSLKIQKLTNNLVIYLTSKLKKPKLVHQSANKLPESITYLELPSNNIT
      PKKLYTYSYKKNLNIQPAHLYMLSTFGHYSYLSIAKKNTITHLIDESTYAP
      LIESPSHPTKLERLYIGNNMLKGIYDHPDLHVPFPYAKKIPNKKNNRFPYAG
      GISINNINLQKTKOISKNDYIFVSORPISDDLKYSIVILNLSISQIKGKIFIK
      LHPKKNNNYVMSLFLNWEINPRLVINEPPELLEPLIYLTNPKGIIGLASSLIYT
      PLISSTOCLSTGELIINLQKYSWMENTMIOEHLIKKFNPTIINDLNGVSNP
      LKTRTPETRLIKSAEFAKSKNYQALPYNNQLASNNITLLGHKALMYNALYNVQK
      IYMERYSDFIYDNTISVDFHSKQKLTWEIKYYSADNRIGDR"

ORIGIN

Query Match 60.0%; Score 27; DB 1; Length 4908;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GCTTTCCTTGATTAAGAATGTTTC 45
    |||||||
DB 1994 GCTTTCCTTGATTAAGAATGTTTC 1968

RESULT 5
AE002098_00
WPCOMMENT
Sequence split into 23 fragments LOCUS AE002098 Accession AE002098
Fragment Name Begin End
AE002098_00 1 110000
AE002098_01 100001 210000
AE002098_02 200001 310000
AE002098_03 300001 410000
AE002098_04 400001 510000
AE002098_05 500001 610000
AE002098_06 600001 710000
AE002098_07 700001 810000
AE002098_08 800001 910000
AE002098_09 900001 1010000
AE002098_10 1000001 1110000
AE002098_11 1100001 1210000
AE002098_12 1200001 1310000
AE002098_13 1300001 1410000
AE002098_14 1400001 1510000
AE002098_15 1500001 1610000
AE002098_16 1600001 1710000
AE002098_17 1700001 1810000
AE002098_18 1800001 1910000
AE002098_19 1900001 2010000
AE002098_20 2000001 2110000
AE002098_21 2100001 2210000
```

```
AE002098_22 2200001 2272360
LOCUS AE002098 2272360 bp DNA circular BCT 26-MAY-2005
DEFINITION Neisseria meningitidis MC58, complete genome.
ACCESSION AE002098 AE002359-AE002564
VERSION AE002098.2 GI:66731897
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis MC58
Neisseria meningitidis MC58
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 2272360)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.V., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Douglass,J.B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Maignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
2 (bases 1 to 2272360)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Ciecko,A.,
Pedon,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Douglass,J.B.A., Mason,T.M., Parksey,D.S., Blair,E., Cittone,H.,
Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.M., Qin,H.,
Vamathevan,J., Gill,J., Scarlato,V., Maignani,V., DeBoy,R.T.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 2272360)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Ciecko,A.,
Pedon,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Douglass,J.B.A., Mason,T.M., Parksey,D.S., Blair,E., Cittone,H.,
Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.M., Qin,H.,
Vamathevan,J., Gill,J., Scarlato,V., Maignani,V., DeBoy,R.T.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Direct Submission
Submitted (18-MAY-2005) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Sequence update by submitter
On or before May 26, 2005 this sequence version replaced
gi:7225225, gi:7225241, gi:7225245, gi:7225252, gi:7225254,
gi:7225269, gi:7225277, gi:7413418, gi:7413419, gi:7225303,
gi:7225314, gi:7225327, gi:7225337, gi:7413420, gi:7413421,
gi:7413423, gi:7225347, gi:7225407, gi:7225416, gi:7225428,
gi:7225435, gi:7225442, gi:7225455, gi:7225470, gi:7225484,
gi:7225498, gi:7225506, gi:7225512, gi:7225523, gi:7225529,
gi:7225537, gi:7225550, gi:7225556, gi:7225573, gi:7225586,
gi:7225599, gi:7413428, gi:7413429, gi:7225631, gi:7225640,
gi:7225650, gi:7225659, gi:7225669, gi:7225683, gi:7225688,
gi:7225697, gi:7413430, gi:7413433, gi:7225726, gi:7225740,
gi:7225757, gi:7225766, gi:7225776, gi:7225783, gi:7225785,
gi:7225809, gi:7225816, gi:7225819, gi:7225845, gi:7225853,
gi:7225863, gi:7225876, gi:7225889, gi:7225898, gi:7225913,
gi:7225923, gi:7225930, gi:7225939, gi:7225948, gi:7225959,
gi:7225969, gi:7225986, gi:7225997, gi:7226012, gi:7226023,
gi:7226036, gi:7226049, gi:7226062, gi:7226072, gi:7226083,
gi:7226100, gi:7226113, gi:7226123, gi:7226133, gi:7226149,
gi:7226162, gi:7226173, gi:7226185, gi:7413445, gi:7413446,
gi:7226218, gi:7226229, gi:7226238, gi:7226246, gi:7226261,
gi:7226273, gi:7226282, gi:7226293, gi:7226304, gi:7226311,
gi:7226320, gi:7226335, gi:7226350, gi:7226363, gi:7413449,
```

gi:7413450, gi:7226401, gi:7413452, gi:7413453, gi:7226436,
gi:7226446, gi:7226457, gi:7226466, gi:7226475, gi:7413455,
gi:7413456, gi:7226457, gi:7413458, gi:7226521, gi:7226533,
gi:7226543, gi:7226555, gi:7226568, gi:7226577, gi:7226589,
gi:7226599, gi:7226609, gi:7413459, gi:7413466, gi:7226631,
gi:7226640, gi:7226651, gi:7226655, gi:7226665, gi:7226674,
gi:7226684, gi:7226690, gi:7226701, gi:7226712, gi:7226724,
gi:7226737, gi:7226744, gi:7226755, gi:7226767, gi:7226775,
gi:7226785, gi:7226794, gi:7226801, gi:7226811, gi:7226820,
gi:7226832, gi:7226844, gi:7226857, gi:7413469, gi:7413470,
gi:7226886, gi:7226894, gi:7226905, gi:7226912, gi:7226928,
gi:7226937, gi:7226947, gi:7226962, gi:7226972, gi:7226980,
gi:7226991, gi:7227004, gi:7413474, gi:7413475, gi:7227034,
gi:7227043, gi:7227054, gi:7227065, gi:7227078, gi:7227087,
gi:7227095, gi:7227103, gi:7227115, gi:7227125, gi:7227136,
gi:7227151, gi:7227159, gi:7227175, gi:7227188, gi:7227203,
gi:7227214, gi:7227229, gi:7227234, gi:7227244, gi:7227249,
gi:7227258, gi:7413476, gi:7413477, gi:7227274, gi:7227292,
gi:7227301, gi:7413479, gi:7413481, gi:7227332, gi:7227345,
gi:7413482, gi:7413484, gi:7227384, gi:7227392, gi:7227405,
gi:7227418, gi:12057208.

FEATURES

source

Location/Qualifiers

1..2272360

/organism="Neisseria meningitidis MC58"

/mol_type="genomic DNA"

/strain="MC58"

/db_xref="taxon:122586"

/note="serogroup: B"

complement(7..498)

/locus_tag="NMB0001"

complement(7..498)

/locus_tag="NMB0001"

complement(7..498)

/note="identified by similarity to EGAD:160712; match to

protein family HMM PF00583"

/codon_start=1

/transl_table=1

/product="putative acetyltransferase"

/protein_id="AAP40480.1"

/db_xref="GI:7225226"

/translation="MNSLFVNDVPTIRLKAGHIGRLVQALPEHMGPEPMSVYKH

AYVRCIKDDELPAFAAYDDSGILGSAVAKHDESPFRITWLGDFVLPETRGK

GIGRLVAVICIGANSLGIKFLYLPDQVIFESHGVVGVGHFHNGSWVTVMRLDV

DVV"

complement(502..897)

/locus_tag="NMB0002"

complement(502..897)

/locus_tag="NMB0002"

/note="identified by Glimmer2; putative"

/codon_start=1

/transl_table=1

/product="hypothetical protein"

/protein_id="AAP40481.1"

/db_xref="GI:7225227"

/translation="MPSDVGIRLQTAFAKMKIKFYELVWVFTSVWAGADBNLL

SISGGRVLAQKNNIDGKIGMOSDLEPARRRLQTAQADIRLEAIPAAVQAQARQ

ABDLQIGVRLAHMAVYAGAGTASGN"

complement(918..2312)

/gene="gltX"

/locus_tag="NMB0003"

complement(918..2312)

/gene="gltX"

/locus_tag="NMB0003"

complement(918..2312)

/gene="gltX"

/locus_tag="NMB0003"

complement(918..2312)

/gene="gltX"

/locus_tag="NMB0003"

complement(918..2312)

/gene="gltX"

BC number="6.1.1.17"
/note="identified by similarity to EGAD:16197; match to
protein family HMM PF00749; match to protein family HMM
TIGR00464"
/codon_start=1
/transl_table=1
/product="glutamate-lyase"
/protein_id="AAP40482.1"
/db_xref="GI:7225228"
/translation="MTVKTRFAPSPTGVIHIGCVRTALFSMAFAHHKGEFLIRIDT
DIARSTAESVNIILDMKRVGLINDADNVVYQTRRFDRYKEVIAELLEKHAAYCYC

Query Match 60.0%; Score 27; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GCTTCCCTTGATTAAGAGTCTTTC 45
DB 78427 GCTTCCCTTGATTAAGAGTCTTTC 78453

RESULT 6
AX043922 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 1 from Patent WO0066791.
ACCESSION AX043922
VERSION AX043922.1 GI:11342850
KEYWORDS

SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.

REFERENCE
AUTHORS
Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,
Maesigani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,
Scarlato, V., Rappuoli, R., Fraser, C. M. and Grandi, G.

TITLE
JOURNAL
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A1 09-NOV-2000;
CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES
source
1..349980
Location/Qualifiers

/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces."
seq 1: 1 to 349980 349980 bases
seq 108: 300001 to 649980 349980 bases
seq 110: 600001 to 949980 349980 bases
seq 111: 900001 to 1249980 349980 bases
seq 112: 120001 to 1549980 349980 bases
seq 113: 1500001 to 1849980 349980 bases
seq 114: 1800001 to 2149980 349980 bases
seq 115: 2100001 to 2273325 173325 bases

ORIGIN

Qy	Db	19	100.0%	Score 27;	DB 6;	Length 349980;
Query Match	Best Local Similarity	100.0%	Pred. No. 11;			
Matches 27;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
RESULT 7	AC094874/c	AC094874*	239048 bp	DNA	linear	HTG 09-MAY-2003
LOCUS	DEFINITION	Rattus norvegicus clone CH230-5N22, *** SEQUENCING IN PROGRESS ***.				
ACCESSION	AC094874	AC094874.7	GI:30466366			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.					
SOURCE	Rattus norvegicus (Norway rat)					
ORGANISM	Rattus norvegicus					
REFERENCE						
AUTHORS						

```

JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 239048)
AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              (bases 1 to 239048)
REFERENCE    Rat Genome Sequencing Consortium.
AUTHORS      Direct Submission
TITLE        Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              On May 9, 2003 this sequence version replaced gi:24818336.
              The sequence in this assembly is a combination of BAC Based reads
              and whole genome shotgun sequencing reads assembled using Atlas
              (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
              in the feature table below represents a scaffold in the Atlas
              assembly (a 'contig-scaffold'). Within each contig-scaffold,
              individual sequence contigs are ordered and oriented, and separated
              by sized gaps filled with Ns to the estimated size. The sequence
              may extend beyond the ends of the clone and there may be sequence
              contigs within a contig-scaffold that consist entirely of whole
              genome shotgun sequence reads. Both end sequences and whole genome
              shotgun sequence only contigs will be indicated in the feature
              table.

---
              Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu

---
              Project Information
              Center project name: GBPK
              Center clone name: CH230-5N22

---
              Summary Statistics
              Assembly program: Atlas;
              Consensus quality: 228277 bases at least Q40
              Consensus quality: 230525 bases at least Q30
              Consensus quality: 232432 bases at least Q20
              Estimated insert size: 234098; sum-of-contigs estimation
              Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

---
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html) .
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 226899: contig of 226899 bp in length
* 226900 226999: gap of unknown length
* 227000 231913: contig of 4914 bp in length
* 231914 232013: gap of unknown length
* 232014 233153: contig of 1140 bp in length
* 233154 233253: gap of unknown length
* 233254 234563: contig of 1310 bp in length
* 234564 234663: gap of unknown length
* 234664 236136: contig of 1473 bp in length
* 236137 236236: gap of unknown length
* 236237 237299: contig of 1063 bp in length
* 237300 237399: gap of unknown length
* 237400 239048: contig of 1649 bp in length.

Location/Qualifiers
1. .239048
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/cclone="CH230-5N22"

misc_feature
1. .1354
/note="wgb_end_extension

```

```

misc_feature      clone_end:Sp6"
                  3320..4109
                  /note="clone boundary
clone_end:Sp6
site:ECORI
end_sequence:BH362997"
misc_feature      complement(219169..220010)
                  /note="clone boundary
                  clone_end:T7
                  site:ECORI
                  end_sequence:BH362996"
misc_feature      225464..226899
                  /note="wgs_end_extension
clone_end:T7"
gap              226900..226999
                  /estimated_length=unknown
misc_feature      227000..228914
                  /note="wgs_end_extension
clone_end:T7"
gap              231914..232013
                  /estimated_length=unknown
gap              233154..233253
                  /estimated_length=unknown
gap              234564..234663
                  /estimated_length=unknown
gap              236137..236236
                  /estimated_length=unknown
gap              237300..237399
                  /estimated_length=unknown

ORIGIN
Query Match      57.3%; Score 25.8; DB 14; Length 239048;
Best Local Similarity 73.3%; Pred. No. 34;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 CGACAGATTCCGCCACCGCTTCTTGATTAAGATTGTTTC 45
Db 93603 CTACAGGATTCCTGCATCCCTTACCTTGTAAAGAAAATGTTTC 93559

RESULT 8
BC016953/c      2603 bp      mRNA      linear      PRI 23-SRP-2002
LOCUS           BC016953
DEFINITION      Homo sapiens, clone IMAGE:3896398, mRNA.
ACCESSION      BC016953.1 GI:23270815
VERSION
KEYWORDS
SOURCE
ORGANISM        Homo sapiens (human)
                 Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                 Homnidae; Homo.
                 1 (bases 1 to 2603)
REFERENCE
AUTHORS         Strausberg,R.
TITLE           Direct Submission
JOURNAL         Submitted (05-NOV-2001) National Institutes of Health, Mammalian
                 Gene Collection (MGC), Cancer Genomics Office, National Cancer
                 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                 USA
REMARK          NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT          Contact: MGC help desk
                 Email: cgapbs-r@mail.nih.gov
                 Tissue Procurement: ATCC
                 cDNA Library Preparation: Life Technologies, Inc.
                 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
                 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
                 Center, Stanford University School of Medicine, Stanford, CA 94305
                 Web site: http://www-shgc.stanford.edu
                 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
                 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
                 R. M.

Clone distribution: MGC clone distribution information can be found

```

```

FEATURES
source
1..2603
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3896398"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH MGC_70"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"

ORIGIN
Query Match      55.6%; Score 25; DB 8; Length 2603;
Best Local Similarity 75.6%; Pred. No. 56;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATTGTT 42
Db 1613 GACAGATTCCGCCCTCCACTGCTTCTCAGTAAGAGTT 1573

RESULT 9
AC145589/c      152506 bp      DNA      linear      ROD 15-MAY-2004
LOCUS           AC145589
DEFINITION      Mus musculus BAC clone RP24-174M10 from chromosome Y, complete
                 sequence.
ACCESSION      AC145589
VERSION        AC145589.3 GI:45120384
KEYWORDS
SOURCE
ORGANISM        Mus musculus (house mouse)
                 Mus musculus
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
                 Sciurognathi; Muridae; Muridae; Murinae; Mus.
                 1 (bases 1 to 152506)
                 Vanbrunt,A., Van Brunt,A., Haglund,K., Meyer,R. and Haakenson,W.
                 The sequence of Mus musculus BAC clone RP24-174M10
                 Unpublished (2001)
REFERENCE
AUTHORS         Wilson,R.K.
TITLE           Direct Submission
JOURNAL         Submitted (19-JUL-2003) Genome Sequencing Center, 4444 Forest Park
                 Parkway, St. Louis, MO 63108, USA
                 3 (bases 1 to 152506)
REFERENCE
AUTHORS         Wilson,R.K.
TITLE           Direct Submission
JOURNAL         Submitted (22-OCT-2003) Genome Sequencing Center, 4444 Forest Park
                 Parkway, St. Louis, MO 63108, USA
                 4 (bases 1 to 152506)
REFERENCE
AUTHORS         Wilson,R.K.
TITLE           Direct Submission
JOURNAL         Submitted (05-MAR-2004) Genome Sequencing Center, 4444 Forest Park
                 Parkway, St. Louis, MO 63108, USA
                 5 (bases 1 to 152506)
REFERENCE
AUTHORS         Wilson,R.K.
TITLE           Direct Submission
JOURNAL         Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
                 Parkway, St. Louis, MO 63108, USA
                 On Mar 5, 2004 this sequence version replaced gi:37806544.
                 Genome Center
                 Center: Washington University Genome Sequencing Center
                 Center code: WUGSC
                 Web site: http://genome.wustl.edu
                 Contact: submissions@watson.wustl.edu
                 Summary Statistics
                 Center project name: M_BB0174M10

```


NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Mouse Chromosome Y Mapping Project (Jessica E. Alfoldi, Helen Skaletsky, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RPCT-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC139328.

FEATURES

source

1.152506
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="Y"
/map="Y"
/clone="RP24-174M10"
/clone_1ib="RPCT-24"
1.43
/note="Sequence derived from one plasmid subclone."
8009.8525
/note="Unresolved tandem repeat."
58030.58040
/note="Sequence derived from one plasmid subclone."
81477.81543
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=16.26 / Sec struct Sc=4.85)"

ORIGIN

Query Match 54.7%; Score 24.6; DB 9; Length 152506;
Best Local Similarity 76.9%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 30; Conservative 0; Mismatches 9;

6 GAATTCGGCCACCGCTTCTGTTGATTAAGATGTTT 44

Db 122822 GAATTCGCCCAATGCTTTCATGTGAACCGAATGTTT 122784

RESULT 10
AC147621/c 166529 bp DNA linear ROD 15-MAY-2004
LOCUS AC147621
DEFINITION Mus musculus BAC clone RP24-315M14 from chromosome Y, complete sequence.
ACCESSION AC147621
VERSION AC147621.3 GI:46358243
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 166529)

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

repeat_region
/rpt_family="ERVK"

VanBrunt, A., Van Brunt, A., Cotton, M. and Bielicki, L.

The sequence of Mus musculus BAC clone RP24-315M14

Unpublished (2001)

2 (bases 1 to 166529)

Wilson, R.K.

Direct Submission

Submitted (22-DEC-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 166529)

Wilson, R.K.

Direct Submission

Submitted (01-JAN-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 166529)

Wilson, R.K.

Direct Submission

Submitted (10-APR-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 166529)

Wilson, R.K.

Direct Submission

Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

On Apr 10, 2004 this sequence version replaced gi:40539137.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0315M14

```
repeat_region 1656.1731 /rpt_family="ERVK"
repeat_region 1803.1902 /rpt_family="ERVK"
repeat_region 3764.3878 /rpt_family="MaLR"
repeat_region 3900.4490 /rpt_family="MaLR"
repeat_region 4489.6799 /rpt_family="L1"
repeat_region 6850.7034 /rpt_family="L1"
repeat_region 7030.7186 /rpt_family="MER1_type"
repeat_region 7187.7312 /rpt_family="MaLR"
repeat_region 7313.7451 /rpt_family="Alu"
repeat_region 8534.8720 /rpt_family="MaLR"
repeat_region 11457.11659 /rpt_family="B2"
repeat_region 14027.14118 /rpt_family="MaLR"
repeat_region 18092.18259 /rpt_family="Alu"
repeat_region 18603.18652 /rpt_family="L1"
repeat_region 18773.18839 /rpt_family="MaLR"
repeat_region 18890.19170 /rpt_family="ERV1"
repeat_region 19462.19733 /rpt_family="MaLR"
repeat_region 20000.20241 /rpt_family="ERVK"
repeat_region 20825.20971 /rpt_family="MaLR"
repeat_region 21706.21842 /rpt_family="Alu"
repeat_region 22418.22549 /rpt_family="Alu"
repeat_region 23262.23397 /rpt_family="Alu"
repeat_region 24072.24220 /rpt_family="Alu"
repeat_region 24833.24967 /rpt_family="Alu"
repeat_region 25543.25694 /rpt_family="Alu"
repeat_region 26172.26436 /rpt_family="Alu"
repeat_region 26530.26668 /rpt_family="L1"
repeat_region 27249.27382 /rpt_family="Alu"
repeat_region 27705.27975 /rpt_family="Alu"
repeat_region 28014.28157 /rpt_family="L1"
repeat_region 28807.28945 /rpt_family="Alu"
repeat_region 29573.29707 /rpt_family="Alu"
repeat_region 30236.30295 /rpt_family="Alu"
repeat_region 30374.30452 /rpt_family="ERV1"
repeat_region 30434.30487 /rpt_family="Alu"
repeat_region 31110.31170 /rpt_family="B4"
repeat_region 31866.31996 /rpt_family="Alu"
```

```
repeat_region 32719.32848 /rpt_family="Alu"
repeat_region 33198.33905 /rpt_family="Alu"
repeat_region 35982.36065 /rpt_family="L1"
repeat_region 36278.37477 /rpt_family="ERVK"
repeat_region 38048.38946 /rpt_family="ERVK"
repeat_region 39943.40253 /rpt_family="ERVK"
repeat_region 40482.40939 /rpt_family="ERV1"
repeat_region 41385.41423 /rpt_family="ERVK"
repeat_region 41424.41534 /rpt_family="U2"
repeat_region 41557.41670 /rpt_family="L1"
repeat_region 42332.42519 /rpt_family="L1"
repeat_region 42520.42578 /rpt_family="L1"
repeat_region 43031.43320 /rpt_family="B4"
repeat_region 43771.43925 /rpt_family="L1"
repeat_region 43932.44030 /rpt_family="B4"
repeat_region 44265.44588 /rpt_family="L1"
repeat_region 45305.45539 /rpt_family="MaLR"
repeat_region 45825.46527 /rpt_family="L1"
repeat_region 46533.47622 /rpt_family="L1"
repeat_region 47688.48296 /rpt_family="L1"
repeat_region 48297.48449 /rpt_family="L1"
```

Query Match 54.7%; Score 24.6; DB 9; Length 166529;
Best Local Similarity 76.9%; Pred. No. 1e+02; Mismatches 9; Indels 0; Gaps 0;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 6 GAATTCGGCCAGCGCTTCTTGATAGATGTTT 44
DB 15649 GGATTCGCCAGCGCTTCTTGATAGATGTTT 15611

RESULT 11
AC107117 3/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AC107117 Accession AC107117
Fragment Name Begin End
AC107117_0 1 110000
AC107117_1 100001 210000
AC107117_2 200001 310000
AC107117_3 300001 359262
Continuation (4 of 4) of AC107117 from base 300001 (AC107117 Rattus norvegicus clone CH;

Query Match 53.8%; Score 24.2; DB 14; Length 59262;
Best Local Similarity 71.1%; Pred. No. 1.4e+02; Mismatches 13; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCACGAAATTCGCCAGCGCTTCTTGATAGATGTTTC 45
DB 48182 CTAGAGGCTTCACCGCGCTTCTTGATAGATGTTTC 48138

RESULT 12

	AC135125	185020 bp	DNA	linear	HTG 20-NOV-2002
LOCUS DEFINITION	Rattus norvegicus clone CH230-265D13, *** SEQUENCING IN PROGRESS				
ACCESSION VERSION KEYWORDS SOURCE ORGANISM	AC135125 AC135125 .2 GI:25138669 HTG, HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.				
REFERENCE AUTHORS	Muzny,D.,Marie,,Metzker,M.,Lee,,Abramzon,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Albrooks,S.,Amiri,A.,Angiano,D.,Anylabechi,V.,Ayagi,A.,Ayodeji,M.,Bacé,E.,Baden,H.,Baladin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Blawie,K.,Blair,U.,Blankenburg,K.,Bylth,P.,Brown,M.,Bryant,N.,Buchay,C.,Burck,P.,Burrell,K.,Calderon,R.,Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,Cleveland,C.,Cookrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,Davila,M.L.,Davis,C.,Davey-Carroll,L.,De Anda,C.,Derdich,D.,Delgado,O.,Denison,S.,Detamo,C.,Ding,Y.,Dirha,K.,Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duval,B.,Evans,K.,Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foister,M.,Foster,P.,Fraser,C.M.,Gabisi,A.,Gall,R.,Garciá,A., Garner T., Garza,M., Gedgeorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,Gunaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C., Hamilton K., Harvey,Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Humm,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B.,Johnson,R.,Jolyvet,A.,Karpatky,S.,Kelly,S., Kelly.S., Khan,Z., King,L., Kovar,C.,Kovits,C., Kraft,C.L., Lebow,H., Leván Z., Lewis L., Li Z., Liu,J., Liu,J., Liu,W., Liu Y., London,P., Longacre,S., Lopez,J., Lorenzenbaev,L., Louiseged,H., Losado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindrathe,M., Mahmoud,M., Malloy,K., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martin.R., Martinez,E., Manthey,S., McLeod,M.P., McNellil,T.Z., Meenen,E., Mlosovajovic,A., Miner,G., Minja,E., Montemayo,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Natr,L., Nakervic,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nowokletemeh,O., Okwonon.G., Olarinpoogon,A., Pal,S., Parks,K., Paternakel,S., Paul,H., Perez,A., Perez,L., Pfankoch.C., Plopper.F., Poindexter.A., Popovic.D., Primus.E., Pu.L.-L., Puzo,M., Quiroz.J., Rachin.E., Reeves.K., Regier,M.A., Reigh.R., Relly,B., Relly,M., Ren.Y., Reuter,M., Richards,S., Riggs.F., Rivers,C., Rodkey.T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders.W., Savery,G., Scherer,S., Scott,G., Shatman.S., Shen.H., Shevty,J., Shivartsbeyan,A., Sisson.I., Slitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Thoror,P., Taylor,C., Taylor.T., Thomas.N., Thomas.S., Tingey,A., Walker,Z., Umami,K., Valias,R., Vera.V., Villasana,D., Waldron,L., Walker.B., Wang,J., Wang,Q., Wang,S., Warren.U., Warren.R., Wei.X., Wiley.F., Williams.G., Willson.R., Wleczyk.R., Wooden.H., Worley,K., Wright,D., Wright.R., Wu.J., Yakub,S., Yen.J., Yoon.L., Yoon.V., Yu.F., Zhang,J., Zhou,U., Zhou.x., Zhao.S., Dunn.D., von Niederhausen,A., Weiss,R., Smithn,D.R., Holt,R.A., Smithn,H.O., Weinstein,G. and Gibbs,R.A.				
TITLE	Unpublished				
JOURNAL REFERENCE	2 (bases 1 to 185020)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 185020)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				

```

JOURNAL
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23507714.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly ('a' contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCZ2
Center clone name: CH230-265D13
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 179552 bases at least Q40
Consensus quality: 181077 bases at least Q30
Consensus quality: 181866 bases at least Q20
Estimated insert size: 181422; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html) .
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      183656: contig of 183656 bp in length
*      183657      183756: gap of unknown length
*      183757      185020: contig of 1264 bp in length.
Location/Qualifiers
    1..185020
     /organism="Rattus norvegicus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10116"
     /clone="CH230-265D13"
        1..1163
         /note="wgs_end_extension
clone_end:T7"
3257..4125
/note="clone_boundary
clone_end:T7
site:-
end_sequence:RXAFC19TVJ"
complement(178207..179102)
/note="clone_boundary
clone_end:Sp6
site:-
end_sequence:RXAFC19TVB"
179376..181250
/note="wgs_end_extension
clone_end:Sp6"
182143..183656
/note="wgs_end_extension
clone_end:Sp6"
183657..183756
/estimated_length=unknown
FEATURES
source
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
gap
ORIGIN

```

Query Match 53.8%; Score 24.2; DB 14; Length 185020;
Best Local Similarity 71.1%; Pred. No. 1.5e+02;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CGACGAAATTCGGCCACCGCTTCTTGATTAAGAAATGTTTC 45
Db 177225 CTAGAGGCTTCACACCGCTACTATGATAAATGACTGTTCC 177181

RESULT 13
LOCUS CQ505083 367 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 36950 from Patent WO0160860.
ACCESSION CQ505083
VERSION CQ505083.1 GI:41471149
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 36950 23-AUG-2001;
FEATURES
Source Millennium Predictive Medicine, Inc. (US)
Location/Qualifiers
1.367
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 53.3%; Score 24; DB 6; Length 367;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAAATGTTT 43
Db 271 CACAAATTAAGCTTCTTCTTCTTGATTAAGAAATGTTT 310

RESULT 14
LOCUS CQ510708 367 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 42575 from Patent WO0160860.
ACCESSION CQ510708
VERSION CQ510708.1 GI:41476972
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 42575 23-AUG-2001;
FEATURES
Source Millennium Predictive Medicine, Inc. (US)
Location/Qualifiers
1.367
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 53.3%; Score 24; DB 6; Length 367;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAAATGTTT 43
Db 271 CACAAATTAAGCTTCTTCTTCTTGATTAAGAAATGTTT 310

RESULT 15
LOCUS AC126059 237515 bp DNA linear HTG 13-NOV-2002
DEFINITION Rattus norvegicus clone CH230-9C16, *** SEQUENCING IN PROGRESS ***
AC126059
AC126059
6 unordered pieces.
AC126059
AC126059.3 GI:24940978
HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guertel, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hune, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Labow, H., Lewis, J., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Loulsegged, H., Lozano, R. J., Lu, X., Ma, T., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelamen, O., Okwou, G., Olariunagoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Ploper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., Pu, L., Pu, M., Quito, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slason, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svetek, A., Tabors, Z., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, P., Umam, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, J., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.
TITLE
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 237515)
AUTHORS Worley, K. C.
DIRECT SUBMISSION

JOURNAL Submitted (03-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237515)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:22772785.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDMJ
Center clone name: CH230-9C16
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226053 bases at least Q40
Consensus quality: 227879 bases at least Q30
Consensus quality: 228965 bases at least Q20
Estimated insert size: 226154; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 220275: contig of 220275 bp in length
* 220276 220375: gap of unknown length
* 220376 232906: contig of 12531 bp in length
* 232907 233006: gap of unknown length
* 233007 234063: contig of 1057 bp in length
* 234064 234163: gap of unknown length
* 234164 235168: contig of 1005 bp in length
* 235169 235268: gap of unknown length
* 235269 236380: contig of 1112 bp in length
* 236381 237515: gap of unknown length
* 237516 237515: contig of 1035 bp in length.
Location/Qualifiers
1. 237515
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9C16"
1. 1341
/note="wgs contig"
220276..220375
/estimated_length=unknown
misc_feature
/note="wgs contig"
220376..221567
misc_feature
/note="wgs contig"
224658..225837
/note="wgs contig"

gap 232907..233006
/estimated_length=unknown
gap 234064..234163
/estimated_length=unknown
gap 235169..235268
/estimated_length=unknown
gap 236381..236480
/estimated_length=unknown
ORIGIN
Query Match 53.3%; Score 24; DB 14; Length 237515;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
By 5 AGAATTCGCCACCGCTTCTTGATTAAGATGTTT 44
DB 82800 AGCAATTCGCCACGATTCCTGTATTAAGCATGCTT 82839
Search completed: April 7, 2006, 18:04:50
Job time : 313.615 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:06 ; Search time 50.7604 Seconds
(without alignments)
5908.370 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45

Sequence: 1 cgacagatccgcacacgc.....ttgtgattaagaatgttcc 45

Scoring table: IDENTITY NTC

Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn190s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	45	2	Aax84283
2	27	60.0	623	3	AAa81833
3	27	60.0	687	3	AAz24472
4	27	60.0	65632	3	AAa81502
5	27	60.0	110000	3	AAa81490
6	27	60.0	349980	3	AAa81490_N_mening
7	24	53.3	367	5	ABV36932
8	24	53.3	367	5	ABV42556
9	23.4	52.0	31	13	ADT77703
10	23.4	52.0	351	6	ABL80063
11	23.4	52.0	3812	13	ADU07328
12	23.4	52.0	4470	12	ADG63001
13	23.2	51.6	720	10	ADH85460
14	22.8	50.7	35	3	AAz24471
15	22.8	50.7	3371	11	ADM02496
16	22.6	50.2	364	3	AAc21971
17	22.4	49.8	546	13	ACN56729
18	22.4	49.8	2576	3	AAz25701
19	22.4	49.8	2576	4	AAH47743

20	22	48.9	6483	6	ABK84534
21	22	48.9	6483	6	ABN97204
22	22	48.9	6483	12	ADH13175
23	22	48.9	6483	14	ABR15062
24	22	48.9	22671	2	AAV31200
25	21.8	48.4	572	4	ABAE2035
26	21.8	48.4	572	4	AAI41973
27	21.8	48.4	572	4	AAK36253
28	21.8	48.4	572	4	AAK10350
29	21.8	48.4	572	4	ABG35934
30	21.8	48.4	1291	4	ABL07613
31	21.8	48.4	1632	8	ADA70091
32	21.6	48.0	3291	4	ABL07612
33	21.6	48.0	585	5	ABV59175
34	21.6	48.0	1099	14	ADM16891
35	21.6	48.0	2688	4	AAI37527
36	21.6	48.0	2688	8	ABK60515
37	21.6	48.0	2688	12	ADJ31265
38	21.6	48.0	2784	3	AAA30848
39	21.6	48.0	2787	6	ABL91241
40	21.6	48.0	2787	8	ACA30627
41	21.6	48.0	2950	3	AAA30847
42	21.6	48.0	3200	2	AAK06816
43	21.6	48.0	11854	2	AAK13243
44	21.6	48.0	11854	6	ABG59038
45	21.6	48.0	37415	12	ADQ97509

ALIGNMENTS

RESULT 1
AAx84283
AAx84283 standard; DNA; 45 BP.

AAx84283;
08-SEP-1999 (first entry)

PCR primer for *Neisseria* CMP-Neu5Ac synthetase coding sequence.
Beta-1,4-galactosyltransferase, IgTB; fusion protein; catalytic domain;
glycosyltransferase; accessory enzyme; nucleotide sugar formation;
saccharide donor; oligosaccharide synthesis; CMP-Neu5Ac synthetase;
carbohydrate structure development; PCR primer; ss.

Synthetic.
Neisseria sp.

WO9931224-A2.

24-JUN-1999.

15-DEC-1998; 98WO-CA001180.

15-DEC-1997; 97US-0069443P.

14-DEC-1998; 98US-00211691.

(CANADA) NAT RES COUNCIL CANADA.

Gilbert M, Young NM, Wakarchuk WM;

WPI, 1999-395174/33.

A new glycosyltransferase fusion protein useful in the enzymatic synthesis of oligosaccharides.

Example 1; Page 40; 63pp; English.

This sequence represents a PCR primer for the *Neisseria* CMP-Neu5Ac synthetase coding sequence. The invention relates to a nucleic acid encoding a fusion protein that comprises a glycosyltransferase catalytic domain and a catalytic domain from an accessory enzyme that is involved

CC in formation of a nucleotide sugar which is a saccharide donor for a
CC glycosyltransferase reaction. The fusion protein is useful in the
CC enzymatic synthesis of oligosaccharides. The fusion proteins are able to
CC catalyze more than one reaction involved in the enzymatic synthesis. This
CC is useful for the development of therapeutic agents that have specific
CC carbohydrate structures. Carbohydrates are involved in recognition
CC elements on the surface of cells. The fusion protein can be used for the
CC synthesis of both natural carbohydrates and synthetic derivatives with
CC novel properties. The fusion polypeptide allows two glycosyltransferase
CC reactions in a single vessel, provides improved yields of end products.
CC Additionally, cleanup and disposal of extra solvents and by-products is
CC reduced. The fusion protein can also use directly different donor
CC analogues and various acceptors with a terminal galactose residue
SQ Sequence 45 BP; 10 A; 12 C; 8 G; 15 T; 0 U; 0 Other;

Query Match 100.0%; Score 45; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGACGAAATTCGCGCACCGCTTCTTGATTAAGAAGTTTTC 45
|||||
Db 1 CGACGAAATTCGCGCACCGCTTCTTGATTAAGAAGTTTTC 45

RESULT 2
AA81833
ID AA81833 standard; DNA; 623 BP.

AC AAA81833;
XX
XX 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_380 SEQ ID NO:380.

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.

XX Neisseria meningitidis.

XX WO200022430-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US023573.

XX 09-OCT-1999; 98US-0103794P.

XX 30-APR-1999; 99US-0132068P.

XX (CHIR) CHIRON CORP.

XX Frazer CM, Hickey B, Peterson J, Tettelin H, Venter JC;
PI Massignani V, Galeotti C, Mora M, Ratti G, Scarbelli M, Scarlato V;
PI Rappelli R, Piza M;

XX WPI; 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.

XX Claim 7; Page 1611; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AA81453 to AA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AA81260 to AA81303 and AA825620 to AA82563 represent Neisseria DNA
CC sequences and their corresponding proteins; AA81254 to AA81259 and
CC AA81304 to AA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AA81322 to AA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid

CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
SQ Sequence 623 BP; 174 A; 118 C; 135 G; 196 T; 0 U; 0 Other;

Query Match 60.0%; Score 27; DB 3; Length 623;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GCTTCCCTTGATTAAGAAGTTTTC 45
|||||
Db 161 GCTTCCCTTGATTAAGAAGTTTTC 187

RESULT 3
AA24472/C
ID AA24472 standard; DNA; 687 BP.

XX AA24472;

XX 17-FEB-2000 (first entry)

XX N. meningitidis siab DNA.

XX CMP sialylate synthetase; siab; cytidine monophosphate; detection;
KW CMP-activated N-acetylneuraminic acid; sialyl acceptor; ds.

XX Neisseria meningitidis.

XX DE19913206-A1.

XX 07-OCT-1999.

XX 24-MAR-1999; 99DE-01013206.

XX 26-MAR-1999; 98DE-01013426.

XX (FESS/) FESSNER W D.

XX Fesener W, Knorst M;

XX WPI; 2000-024923/03.

XX P-PSDB; AAY50798.

XX Production of CMP-sialate synthetase protein useful for producing CMP-
PT activated N-acetylneuraminic acid analogs and non-naturally sialylated
PT oligosaccharides and glycoconjugates.

XX Claim 5; Page 19-20; 26pp; German.

XX This invention describes a novel method for the production of a CMP
CC (cytidine monophosphate)-sialate synthetase protein which comprises
CC culturing a prokaryotic host organism transformed with an expression
CC vector containing a structural gene for the protein, where the start
CC codon of the structural gene is 8-12 bases downstream from a ribosome
CC binding site (RBS). The protein is useful for producing CMP-activated N-
CC acetylneuraminic acid analogs by reacting the analogs with cytidine
CC triphosphate in the presence of the protein and for producing non-
CC naturally sialylated oligosaccharides and glycoconjugates, which are of

CC therapeutic interest because of their sialidase resistance. The protein
CC is also useful for detecting sialyl acceptors by reacting the acceptor
CC with a fluorescent sialic acid derivative in the presence of the protein.
CC cytidine triphosphate and a sialyl transferase and detecting the
CC fluorescently modified acceptor by selective chromatography. This
CC sequence encodes a CMP (cytidine monophosphate)-sialate synthetase
CC protein, sialB, isolated from *Neisseria meningitidis* serogroup B
XX
SQ Sequence 687 BP; 218 A; 154 C; 122 G; 193 T; 0 U; 0 Other;
Query Match 60.0%; Score 27; DB 3; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GCTTTCCTTGATTAAGATGTTTC 45
DB 684 GCTTTCCTTGATTAAGATGTTTC 658
RESULT 4
ID AAA81502 standard; DNA; 65632 BP.
XX
AC AAA81502;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_49 SEQ ID NO:49.
XX
KM *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM *Meningococcus B*; MenB; ds.
XX
OS *Neisseria meningitidis*.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masihami V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be used
PT in the diagnosis and treatment of *N. meningitidis* infection and other
PT Neisserial infections, for example, *N. gonorrhoea*.
XX
PS Claim 7; Page 1331-1350; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
CC from *Neisseria* genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed *Neisseria meningitidis* genomic DNA sequences,
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent *Neisseria* DNA
CC sequences and their corresponding proteins, AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC *Neisseria meningitidis* DNA sequences; and AAA81322 to AAA81452 represent
CC *Neisseria meningitidis* MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to *Neisseria* bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against *Meningococcus B*, against all serotypes; and/or against all

CC pathogenic *Neisseriae*. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious *Meningococcus B*
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 65632 BP; 16704 A; 16327 C; 15474 G; 17126 T; 0 U; 1 Other;
Query Match 60.0%; Score 27; DB 3; Length 65632;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GCTTTCCTTGATTAAGATGTTTC 45
DB 55738 GCTTTCCTTGATTAAGATGTTTC 55712
RESULT 5
ID AAA81490_00
WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490
WP Fragment Name Begin End
WP AAA81490_00 1 110000
WP AAA81490_01 100001 210000
WP AAA81490_02 200001 310000
WP AAA81490_03 300001 410000
WP AAA81490_04 400001 510000
WP AAA81490_05 500001 610000
WP AAA81490_06 600001 710000
WP AAA81490_07 700001 810000
WP AAA81490_08 800001 910000
WP AAA81490_09 900001 1010000
WP AAA81490_10 1000001 1110000
WP AAA81490_11 1100001 1210000
WP AAA81490_12 1200001 1310000
WP AAA81490_13 1300001 1410000
WP AAA81490_14 1400001 1437668
ID AAA81490 standard; DNA; 1437668 BP.
XX
AC AAA81490;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
XX
KM *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM *Meningococcus B*; MenB; ds.
XX
OS *Neisseria meningitidis*.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masihami V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be used

PT in the diagnosis and treatment of *N. meningitidis* infection and other
PT Neisserial infections, for example, *N.gonorrhoea*.
XX
XX
PS Claim 7; Page 866-1272; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAB1453 to AAB2414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAB1260 to AAB1303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAB1254 to AAB1259 and
CC AAB1304 to AAB1321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAB1322 to AAB1452 represent
CC Neisseria meningitidis Mena polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against *Meningococcus B*; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious *Meningococcus B*
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 1437668 BP; 344338A; 353206C; 385074G; 355045T; 0U; 50ther;
Query Match 60.0%; Score 27; DB 3; Length 110000;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GCTTTCCTTGATTAAGAATGTTTC 45
DB 78420 GCTTTCCTTGATTAAGAATGTTTC 78446
RESULT 6
AAB21544
ID AAB21544 standard; DNA; 349980 BP.
XX
XX AAB21544;
AC
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KM ds.
XX
OS Neisseria meningitidis.
XX
XX WO200066791-A1.
PN
XX
PD 09-NOV-2000.
XX
XX 08-MAR-2000; 2000MO-US005928.
PF
XX
XX 30-APR-1999; 199US-0132068P.
PR 08-OCT-1999; 99MO-US023573.
PR 28-FEB-2000; 2000GB-00004695.
XX
XX (CHIR) CHIRON CORP.
PA
XX (GENO-) INST GENOMIC RES.
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarcellini M, Scarlato V;

PI Rappuoli R, Frazer CM, Grandi G;
XX
XX WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAB21544 and AAB21607 to AAB21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAB21544 is repeated at the beginning of
CC AAB21607, the last 49980 bp of AAB21607 are repeated at the beginning of
CC AAB21608, and so on). AAB21545 to AAB21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAB21589 to AAB21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which bind to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used
XX
SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 U; 0 Other;
Query Match 60.0%; Score 27; DB 3; Length 349980;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GCTTTCCTTGATTAAGAATGTTTC 45
DB 78420 GCTTTCCTTGATTAAGAATGTTTC 78446
RESULT 7
ABV36932
ID ABV36932 standard; CDNA; 367 BP.
XX
XX ABV36932;
AC
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker CDNA 36923.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW human; prostatic cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
PN
XX
PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001MO-US005171.
PF
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0235281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JB;

XX MPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 7608; 11750pp; English.
XX
The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SO Sequence 367 BP; 88 A; 82 C; 98 G; 97 T; 0 U; 2 Other;

Query Match 53.3%; Score 24; DB 5; Length 367;
Best Local Similarity 75.0%; Pred. No. 8.1;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTTT 43
DB 271 CACAAATTAAGCTCTACTTCTTCTTGATTAAGAATGTTT 310

RESULT 8
ABV42556
ID ABV42556 standard; cDNA; 367 BP.
XX
AC ABV42556;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 42547.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2_g.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PRDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JR;
XX
XX MPI; 2001-662795/76.
XX
Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8517; 11750pp; English.
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SO Sequence 367 BP; 88 A; 82 C; 98 G; 97 T; 0 U; 2 Other;

Query Match 53.3%; Score 24; DB 5; Length 367;
Best Local Similarity 75.0%; Pred. No. 8.1;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTTT 43
DB 271 CACAAATTAAGCTCTACTTCTTCTTGATTAAGAATGTTT 310

RESULT 9
ADT77703
ID ADT77703 standard; DNA; 31 BP.
XX
AC ADT77703;
XX
DT 13-JAN-2005 (first entry)
XX
DB Neisseria meningitidis capsule synx gene PCR primer capRI.
XX
XX Vaccine; attenuation; synx gene; antibacterial; PCR; primer; ss.
XX
XX Neisseria meningitidis.
XX
PN WO2004089408-A2.
XX
PD 21-OCT-2004.
XX
PF 07-APR-2004; 2004WO-GB001504.
XX
PR 07-APR-2003; 2003GB-00008691.
XX
PA (XENO-) XENOVA RES LTD.
XX
PI McClean CS, Keen SW, Martin GM;
XX
XX MPI; 2004-748603/73.
XX

XX
XX New attenuated mutant Neisseria bacterium comprising an auxotrophic
PT attenuating mutation, a capsule mutation and a mutation which reduces
PT bacterial recombination or exogenous DNA uptake, used as vaccine against
PT Neisseria infection.
XX
PS Example; SEQ ID NO 9; 40pp; English.
XX
XX The present sequence is that of PCR primer capRI for the Neisseria
CC meningitidis synx gene, which is involved in bacterial capsule formation.
CC PCR amplification of strain B16B6 genomic DNA using capRI ADT77702 and
CC capRI generated a 3.1 kb fragment extending from the 5' end of the slab
CC gene to midway through the ctrA gene. This PCR product was used as
CC template for the PCR amplification of left and right hand flanking
CC regions of the synx gene. Primer capRI was used with primer capdelFI
CC ADT77705 to generate a 668 bp right hand flanking region. This was used
CC in the generation of plasmids for synx deletion and hence in the
CC construction of an arCB/synx/reca triple mutant live attenuated vaccine
CC strain. The invention provides an attenuated strain of a Neisseria
CC bacterium, including N. meningitidis, that retains immunogenicity and
CC ability to grow in culture, without causing neisserial disease in
CC subjects to which it is administered as a vaccine. The attenuated

CC Neisseria bacterium, or a live or killed preparation of it, has the
 CC following mutations: (a) an auxotrophic attenuation mutation, e.g. an
 CC ArcA or an ArcB mutation; (b) a capsule mutation that affects capsule
 CC integrity and/or causes the capsule to be of reduced thickness or absent,
 CC e.g. a mutation in the *synX* gene or the *gals* gene; and (c) a mutation
 CC which reduces bacterial recombination or exogenous DNA uptake, e.g. a
 CC mutation in a bacterial recombinase gene such as *recA*, or alternatively a
 CC capsid mutation. A claimed pharmaceutical composition for use as a vaccine
 CC comprises the live or killed mutant *Neisseria* bacterium.

XX Sequence 31 BP; 8 A; 5 C; 6 G; 12 T; 0 U; 0 Other;

Query Match 52.4%; Score 23.6; DB 13; Length 31;

Best Local Similarity 86.7%; Pred. No. 6.7;

Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 10 TCCGCGACGCGCTTCTTGATTAAGAT 39
 2 TCCGATATCGCTTCTTGATTAAGAT 31

RESULT 10
 ABL80063/C
 ID ABL80063 standard; cDNA; 351 BP.

XX ABL80063;
 XX 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:3041.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX MO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX MPI; 2002-122075/16.

PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide.

XX Claim 1; SEQ ID NO 3041; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumor
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI)
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
 CC or antigen presenting cells that express (II). (I) has cytostatic
 CC activity. An oligonucleotide (IV) that hybridizes to (SI) can be used for
 CC detecting ovarian cancer in a patient's biological sample preferably
 CC serum or ovarian tissue. The method comprises contacting a biological
 CC sample from a patient with (IV), detecting the amount of polynucleotide
 CC hybridizing to (IV) and comparing the amount to a predetermined cutoff
 CC value and thereby detecting ovarian cancer in the patient, where the
 CC amount of polynucleotide hybridizing to (IV) is detected preferably by
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
 CC useful for stimulating and/or expanding T cells specific for an ovarian
 CC tumor protein comprising contacting T cells with (III) or (II). (III) is
 CC useful in design and preparation of ribozyme molecules for inhibiting

CC expression of the tumour polypeptides and proteins in tumour cells; and
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
 CC library using well known techniques

XX Sequence 351 BP; 74 A; 106 C; 91 G; 80 T; 0 U; 0 Other;

Query Match 52.0%; Score 23.4; DB 6; Length 351;

Best Local Similarity 73.2%; Pred. No. 14;

Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 2 GACAGATTCCGCGACGCGCTTCTTGATTAAGATGTT 42
 240 GACAGATTCCGCGCGCTTCTTGATTAAGATGTT 200

RESULT 11
 ADR07328/C
 ID ADR07328 standard; cDNA; 3812 BP.

XX ADR07328;
 XX 04-NOV-2004 (first entry)

DE Full length human cDNA useful for treating neurological disease Seq 834.
 XX gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
 XX osteoporosis; neurological disease; Alzheimer's disease;
 XX Parkinson's disease; dementia; short memory; cancer;
 XX sense or motor function; emotional reaction; fear response; panic;
 XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 XX tranquiliser.

XX Homo sapiens.

XX EPI447413-A2.

XX 18-AUG-2004.

XX 12-FEB-2004; 2004EP-00003145.

XX 14-FEB-2003; 2003JP-00102207.

XX 09-MAY-2003; 2003JP-00131452.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

XX Wakamatsu A, Ishii S, Nagai K, Irie R;

XX MPI; 2004-583265/57.

XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 834; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNA. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cytostatic and tranquiliser activities. This polynucleotide is a full
 CC length human cDNA sequence of the invention. NOTE: This sequence is not
 CC given in the sequence listing of the specification but can be obtained on

CC CD-ROM from the European Patent Office, Vienna Sub-office.

XX Sequence 3812 bp; 778 A; 1002 C; 1108 G; 924 T; 0 U; 0 Other;

XX SQ

Query Match 52.0%; Score 23.4; DB 13; Length 3812;

Best Local Similarity 73.2%; Pred. No. 25;

Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACAGAAATTCGCGCCACCGCTTCTTGATTAAGATGTT 42

DB 2857 GACACGATTCGCGCTCCACATGCTCTCTCAGTAACAAGCTT 2817

RESULT 12

ADQ63001/c

ID ADQ63001 standard; cDNA; 4470 BP.

XX ADQ63001;

XX 07-OCT-2004 (first entry)

XX Novel human cDNA sequence #162.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytosolic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.

OS Homo sapiens.

XX EP1440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Tsogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX MPI: 2004-535376/52.

XX P-PSDB; ADQ65189.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 162; 2449bp; English.

XX The invention relates to 2495 novel polynucleotides (1) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.

XX Sequence 4470 BP; 909 A; 1191 C; 1285 G; 1085 T; 0 U; 0 Other;

XX Query Match 52.0%; Score 23.4; DB 12; Length 4470;

XX Best Local Similarity 73.2%; Pred. No. 26;

XX Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACAGAAATTCGCGCCACCGCTTCTTGATTAAGATGTT 42

DB 3512 GACAGAAATTCGCGCTCCACATGCTCTCTCAGTAACAAGCTT 3472

RESULT 13

ADH85460

ID ADH85460 standard; DNA; 720 BP.

XX ADH85460;

XX 22-APR-2004 (first entry)

XX Enterococcus faecalis polynucleotide #3345.

XX Enterococcus faecalis infection; transcription regulatory element;

XX antibacterial; gene; ds.

XX Enterococcus faecalis.

XX US6617156-B1.

XX 09-SEP-2003.

XX 13-AUG-1998; 98US-00134000.

XX 15-AUG-1997; 97US-0055778P.

XX (DOUC/) DOUCETTE-STAMM L A.

XX (BUSH/) BUSH D.

XX Doucette-Stamm LA, Bush D;

XX MPI: 2003-895394/82.

XX P-PSDB; ADH88865.

XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis

XX polypeptide, useful for preparing a composition for diagnosing or

XX treating E. faecalis infection.

XX Disclosure; SEQ ID NO 3345; 193bp; English.

XX The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polynucleotide of the invention.

XX Sequence 720 BP; 244 A; 93 C; 147 G; 236 T; 0 U; 0 Other;

XX Query Match 51.6%; Score 23.2; DB 10; Length 720;

XX Best Local Similarity 77.8%; Pred. No. 21;

XX Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GACAGAAATTCGCGCCACCGCTTCTTGATTAAGA 37

DB 36 GAAAGAAATTTGCGATTCCTTCTTGATTAAGA 71

RESULT 14

AAZ24471

ID AAZ24471 standard; DNA; 35 BP.

XX AAZ24471;

XX 17-FEB-2000 (first entry)

XX N. meningitidis siab PCR primer 2.

XX CMP sialylate synthetase; siab; cytidine monophosphate; detection;

XX CMP-activated N-acetylneuraminic acid; sialyl acceptor; PCR primer; ss.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:00:12 ; Search time 353.403 Seconds
(without alignments)
5957.564 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45
Sequence: 1 cgacagatccgcacacgc.....ttcgtacgaagtgtttc 45

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41080325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_hnc:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_est7:
9: gb_gsa1:
10: gb_gsa2:
11: gb_gsa3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24.6	54.7	672	9	AZ485731
2	24.6	54.7	1958	4	AK017055 Mus muscu
3	24.4	54.2	625	9	BH333088 CH230-59K
4	24.4	54.2	833	2	BE641022 Cr12.2 A2
5	24.4	54.2	1022	5	BUI58055 AGENCOURT
6	23.8	52.9	283	10	CE688260 tigr-gss-
7	23.8	52.9	500	3	BP188097 BP188097
8	23.8	52.9	815	5	BK698840 BK698840
9	23.6	52.4	587	9	BH333091 CH230-59K
10	23.6	52.4	780	9	CC929966 ZMMBB053
11	23.6	52.4	888	4	CNS09115 Single re
12	23.6	52.4	962	10	CL994393 ZMMBB000
13	23.6	52.4	1087	10	CW001860 ZMMBB000
14	23.4	52.0	284	1	BB159100 BB159100
15	23.4	52.0	351	1	AA292589 zc31f01.x
16	23.4	52.0	416	3	BM694767 U1-B-C11-
17	23.4	52.0	433	7	CO997065 pm01-15m
18	23.4	52.0	540	10	CG603013 OST277031
19	23.4	52.0	501	8	DN648303 GT023.62
20	23.4	52.0	576	7	CV002745 pm01-19m
21	23.4	52.0	592	7	CK534070 rtwgpb0.00
22	23.4	52.0	634	10	AG095231 Pan tlogl

23	23.4	52.0	653	8	CX185987	CX185987 B01.45-76
24	23.4	52.0	670	6	CA924883	CA924883 MTU7TU.P1
25	23.4	52.0	674	2	BC911192	BC911192 602807115
26	23.4	52.0	754	4	DN054816	DN054816 UGI_CABAB
27	23.4	52.0	2035	4	HSW807446	BX647302 Homo sapi
28	23.2	51.6	407	5	BO475193	BO475193 ciciendela
29	23.2	51.6	430	1	AM659916	AM659916 97576 MAR
30	23.2	51.6	554	8	DR872146	DR872146 JGI CABG1
31	23.2	51.6	566	8	B2760599	B2760599 623_302.B
32	23.2	51.6	741	9	BH498370	BH498370 BCGY30TF
33	23.2	51.6	750	10	CG852499	CG852499 ZMMBB034
34	23	51.1	260	1	AV360233	AV360233 AV360233
35	23	51.1	393	9	AO168424	AO168424 HS_3151.B
36	23	51.1	533	6	CB422282	CB422282 595372 MA
37	23	51.1	539	6	CA968909	CA968909 CCL03a10
38	23	51.1	559	6	CB421928	CB421928 594988 MA
39	23	51.1	612	8	DT110051	DT110051 JGI ANNO6
40	23	51.1	612	8	DT115631	DT115631 JGI ANNO3
41	23	51.1	676	8	CB429698	CB429698 605532 MA
42	23	51.1	679	8	DN543254	DN543254 1392090 M
43	23	51.1	680	8	DN541153	DN541153 1386690 M
44	23	51.1	729	7	CK247530	CK247530 EST731167
45	23	51.1	729	11	CR312301	CR312301 mte1-34P2

ALIGNMENTS

RESULT 1
AZ485731
LOCUS
DEFINITION
1M0313018F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0313018 F, genomic survey sequence.

ACCESSION
AZ485731
VERSION
AZ485731.1 GI:10651816
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
GSS.

REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Ham1,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingley,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
1 (bases 1 to 672)

AUTHORS

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)

JOURNAL

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., StG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

TITLE

Email: ddunn@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Plate: 0313 row: 0 column: 18
Seq primer: CCGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 672.

FEATURES

source
1..672
location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0313018"
/sex="Male"
/lab_host="E. Coli strain XU10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PMD42nv, Purified genomic DNA from M."

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Robert Strausberg, Ph.D.
EMAIL cgaabs-r@mail.nih.gov

```
/organism="Canis familiaris"  
/mol_type="genomic DNA"  
/strain="Standard Poodle"  
/db_xref="taxon:9615"  
/clone_ltb="Dog Library"
```

ACCESSION	BX698840
VERSION	BX698840.1
KEYWORDS	EST.
SOURCE	Xenopus tropicalis (western clawed frog)
ORGANISM	Xenopus tropicalis

Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;

REFERENCE
1 (bases 1 to 815)
Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M., and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
TITLE
JOURNAL
COMMENT
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu085e03.q1kT7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..815
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TNeu085e03"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Query Match 52.4%; Score 23.8; DB 5; Length 815;
Best Local Similarity 72.1%; Pred. No. 2.3e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 3 ACAGATTCCGCGACCGCTTCTTGATTAAGATGTTTC 45
Db 312 AAAAATTCAAGACACTTATCTTGCTGAAATCTTTTC 354

RESULT 9
BH333091 587 bp DNA linear GSS 03-DEC-2001
LOCUS CH230-59K12.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION BH333091
ACCESSION BH333091
VERSION BH333091.1 GI:17263805
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 587)
Zhao, S., Shetty, J., Shateman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregiorgis, B., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-59K12.TV
Contact: Shaying Zhao
Department of Bukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).

TITLE
JOURNAL
COMMENT

Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 59 row: K column: 12
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..587
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNhd/MCW"
/db_xref="taxon:10116"
/clone="CH230-59K12"
/sex="female"
/cell_type="Brain"
/clone_11b="CHORI-230 Segment 1"
/note="Vector: pTARBA2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SsNhd/MCW) BAC library produced by
Pieter de Jong"

FEATURES
source
1..587
Location/Qualifiers

ORIGIN
Query Match 52.4%; Score 23.6; DB 9; Length 587;
Best Local Similarity 76.3%; Pred. No. 2.6e+02;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 3 ACAGATTCCGCGACCGCTTCTTGATTAAGATG 40
Db 39 ACAGATCCGCTACCTTACCTTGATTAACAGTG 76

RESULT 10
CC929966 780 bp DNA linear GSS 11-AUG-2003
LOCUS ZMMBc0538U06r ZMMBc Zea mays genomic clone ZMMBc-0538U06 3'
DEFINITION CC929966
ACCESSION CC929966
VERSION CC929966.1 GI:33581064
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 780)
Bharti, A.K., Young, S., Kavchok, S., Kaizer, G., Bronzino, A.C.,
Rouzaud, K., Fuks, G., Yu, Y., Wang, R. and Messing, J.
Sequencing of the maize genome at PGIR (2003b)
Unpublished (2003)
Contact: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 510.
Location/Qualifiers
1..780
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBc0538U06"
/lab_host="E. coli DH10B"
/clone_11b="ZMMBc"
/note="Vector: pTARBA2.1; Site_1: BamHI; Site_2: BamHI"

ORIGIN
Query Match 52.4%; Score 23.6; DB 9; Length 780;
Best Local Similarity 76.3%; Pred. No. 2.7e+02;

Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 8 ATTCCGACCGCTTCTTGATTAGATTGTTTC 45
 |||||
 445 ATTCGCCAATGGCTTCTTGATTAGATTGTTTC 482

RESULT 11
 CNS09L15/ 888 bp mRNA linear HTC 22-APR-2005

LOCUS Single read from an extremity of a full-length cDNA made from
 DEFINITION Anopheles gambiae total adult females. 3-PRIME end of clone
 PK0AAC46A06 of strain 6-9 of Anopheles gambiae (African malaria
 mosquito).

ACCESSION BX063945.1 GI:27637226

KEYWORDS HTc.
 SOURCE Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae

REFERENCE 1
 AUTHORS Gomez,S.M., Egliseier,K., Segurens,B., Dehoux,P., Couloux,A.,
 Scarpelli,C., Wincker,P., Weissenbach,J., Brey,P.T. and Roth,C.W.
 TITLE Pilot Anopheles gambiae full-length cDNA study: sequencing and
 initial characterization of 35,575 clones
 JOURNAL Genome Biol. 6 (4), R39 (2005)
 PUBMED 1583126
 REFERENCE 2 (bases 1 to 888)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

FEATURES
 source
 1..888
 Location/Qualifiers
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="6-9"
 /db_xref="taxon:7165"
 /clone="PK0AAC46A06"
 /plasmid="PME18S-FL"
 /note="end : 3-PRIME"

ORIGIN

Query Match 52.4%; Score 23.6; DB 4; Length 888;
 Best Local Similarity 76.3%; Pred. No. 2.8e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 3 ACAGATTCCGACCGCTTCTTGATTAGATTG 40
 |||||
 581 ACCGAGCGCGCACCGCTTCTTGATTAGATTG 544

RESULT 12
 CL994393 962 bp DNA linear GSS 23-SEP-2004
 LOCUS ZMMBHf0006h02.r ZMMBHf Zea mays genomic clone ZMMBHf0006h02 3',
 DEFINITION genomic survey sequence.
 ACCESSION CL994393
 VERSION CL994393.1 GI:52562471
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE 1
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 962)
 Ma,J., Samiguel,P., Liu,R., Haller,K., Soderlund,C. and
 Bennetzen,J.
 TITLE ZMMBH sequences
 JOURNAL Unpublished (2004)

COMMENT

Contact: Jeff Bennetzen
 Bennetzen Lab
 The University of Georgia
 Department of Genetics, C426a Life Sciences Building, Athens, GA
 30602, USA
 Tel: 706-542-3698
 Fax: 706-583-0972
 Email: maize@uga.edu
 Plate: 0006 row: h column: 02
 Class: BAC ends.

FEATURES
 source

1..962
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBHf0006h02"
 /tissue_type="immature ear"
 /dev_stage="6-8 weeks"
 /lab_host="DH10B"
 /clone_lib="ZMMBHf"
 /note="Vector: TOPOpcr4; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN

Query Match 52.4%; Score 23.6; DB 10; Length 962;
 Best Local Similarity 76.3%; Pred. No. 2.8e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 3 ACAGATTCCGACCGCTTCTTGATTAGATTG 40
 |||||
 913 ACAGATTCTCTTCTTGATTAGATTACAGG 950

RESULT 13
 LOCUS ZMMBHg0002h24.f ZMMBHg Zea mays genomic clone ZMMBHg0002h24 5',
 DEFINITION genomic survey sequence.
 ACCESSION ZMMBHg0002h24.f
 VERSION ZMMBHg0002h24.f
 KEYWORDS GSS.
 SOURCE Zea mays

ORGANISM Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 1087)
 Ma,J., Samiguel,P., Liu,R., Haller,K., Soderlund,C. and
 Bennetzen,J.

REFERENCE 1
 AUTHORS Zea mays
 Bennetzen,J.

TITLE ZMMBH sequences
 JOURNAL Unpublished (2004)
 COMMENT Contact: Jeff Bennetzen
 Bennetzen Lab
 The University of Georgia
 Department of Genetics, C426a Life Sciences Building, Athens, GA
 30602, USA
 Tel: 706-542-3698
 Fax: 706-583-0972
 Email: maize@uga.edu
 Plate: 0002 row: h column: 24
 Class: BAC ends.

FEATURES
 source
 1..1087
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBHg0002h24"
 /tissue_type="immature ear"
 /dev_stage="6-8 weeks"
 /lab_host="DH10B"
 /clone_lib="ZMMBHg"
 /note="Vector: TOPOpcr4; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN

Query Match 52.4%; Score 23.6; DB 10; Length 1087;
 Best Local Similarity 76.3%; Pred. No. 2.9e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 AGAATTCGCGCACCGCTTCTTGATTAAGAATG 40
 |||||
 Db 913 AGAATTCCTCTTGCTTCTTGATTAAGAATG 950
 *:

RESULT 14 284 bp mRNA linear EST 29-JUN-2000
 BBI59100 BBI59100 RIKEN full-length enriched, 16 days neonate thymus Mus
 LOCUS musculus cDNA clone A130043N04 3', mRNA sequence.
 DEFINITION
 ACCESSION BBI59100 GI:8815030
 VERSION BBI59100.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE

AUTHORS

Komono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
 Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
 Hirokane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
 Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
 Kusabata, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
 Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
 Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Sugihara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,
 Takahashi, F., Tomimaga, N., Toyata, T., Tsunoda, Y., Watabiki, A.,
 Watanabe, S., Yamamura, T., Yamazaki, I., Yano, R., Yasunishi, A.,
 Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and
 Hayashizaki, Y.

RIKEN Mouse ESTs (Komono, H., et al.)
 Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
 Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 The most stable and thermosensitive of thermolabile enzymes by
 thermal stabilization and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsuina, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Osawa, Y., Muramatsu, M.,
 Okazaki, Y., and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers
 1..284
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="A130043N04"
 /tissue_type="thymus"
 /dev_stage="16 days neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 16 days neonate"

ORIGIN

Query Match 52.0%; Score 23.4; DB 1; Length 284;
 Best Local Similarity 73.2%; Pred. No. 2.8e+02;
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 5 AGAATTCGCGCACCGCTTCTTGATTAAGAATGTTTC 45
 |||||
 Db 115 AGCTTCCCGCACGATTTCTGATTATTAAGAATGTTTCC 155

RESULT 15 351 bp mRNA linear EST 08-AUG-1997
 LOCUS AA292589
 DEFINITION zt31f01.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
 IMAGE:723961 5', mRNA sequence.
 ACCESSION AA292589
 VERSION AA292589.1 GI:1940583
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 351)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
 Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacey, M., Le, M., Le, N.,
 Marais, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Scheinberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
 and Watters, M.

Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LNL, contact the
 IMAGS Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1162 Std Error: 0.00
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 340.
 Location/Qualifiers
 1..351

FEATURES

source

Location/Qualifiers
 1..351
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:723961"
 /sex="Female"
 /tissue_type="ovarian tumor"
 /lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares ovary tumor NbHOT"
 /note="Organ: ovary; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer (5'
 TGTACCAATCTGAGAGGAGCGCGCGGCTTTTCTTTTCTTTT 3').
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo. "

ORIGIN

Query Match 52.0%; Score 23.4; DB 1; Length 351;
 Best Local Similarity 73.2%; Pred. No. 2.9e+02;
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGAATGTT 42
 |||||
 DB 240 GACAGATTCCGCCCTCCACTGTCCTTCTCAGTAACAGCTT 200

Search completed: April 7, 2006, 18:17:07
 Job time : 360.403 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:17:47 ; Search time 16.0996 Seconds
(without alignments)
4968.475 Million cell updates/sec

Title: US-09-211-691-4
Perfect score: 45.4
Sequence: 1 cgacagaattccgccaccgc.....ttgtgataagaattttcc 45

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/1-COMB.seq:*
2: /cgn2_6/prodata/1/ina/5-COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A-COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B-COMB.seq:*
5: /cgn2_6/prodata/1/ina/8-COMB.seq:*
6: /cgn2_6/prodata/1/ina/PC/US-COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP-COMB.seq:*
8: /cgn2_6/prodata/1/ina/RB-COMB.seq:*
9: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.2	51.6	720	3	US-09-134-000C-3345
2	23	51.1	601	3	US-09-949-016-156329
3	23	51.1	101300	3	US-09-949-016-156329
4	22.6	50.2	364	3	US-09-513-999C-26046
5	22.4	49.8	2576	3	US-09-265-108-1
6	22.4	49.8	2576	3	US-09-479-264-1
7	22	48.9	22671	3	US-08-976-259-14
8	22	48.9	22671	3	US-09-956-004-14
9	21.8	48.4	10554	3	US-09-949-016-12320
10	21.8	48.4	10554	3	US-09-949-016-16084
11	21.6	48.0	51291	3	US-09-949-016-15362
12	21.6	48.0	1230230	3	US-09-198-452A-1
13	21.6	48.0	1230230	3	US-09-438-185A-1
14	21.4	47.6	2403	3	US-09-489-039A-70
15	21.4	47.6	4266	3	US-09-489-039A-7
16	21.2	47.1	601	3	US-09-949-016-72448
17	21.2	47.1	819	3	US-09-601-198-45
18	21.2	47.1	132871	3	US-09-949-016-13863
19	21.2	47.1	161607	3	US-09-949-016-12210
20	21.2	47.1	670689	3	US-09-949-016-12505
21	21.2	47.1	670689	3	US-09-949-016-14207
22	21	46.7	2749	3	US-09-619-353-11
23	21	46.7	2923	3	US-09-619-353-9
24	20.8	46.2	381	3	US-09-328-352-1745

C 25	20.8	46.2	966	2	US-08-181-271A-24	Sequence 24, Appl
C 26	20.8	46.2	966	2	US-08-449-315-24	Sequence 24, Appl
C 27	20.8	46.2	966	2	US-08-444-803-24	Sequence 24, Appl
C 28	20.8	46.2	966	2	US-08-449-043-24	Sequence 24, Appl
C 29	20.8	46.2	966	2	US-08-456-263A-24	Sequence 24, Appl
C 30	20.8	46.2	966	2	US-08-455-416-24	Sequence 24, Appl
C 31	20.8	46.2	966	2	US-08-455-244-24	Sequence 24, Appl
C 32	20.8	46.2	966	2	US-08-454-876-24	Sequence 24, Appl
C 33	20.8	46.2	966	2	US-08-457-364-24	Sequence 24, Appl
C 34	20.8	46.2	966	2	US-08-456-262-24	Sequence 24, Appl
C 35	20.8	46.2	966	2	US-08-456-240-24	Sequence 24, Appl
C 36	20.8	46.2	966	2	US-08-455-736-24	Sequence 24, Appl
C 37	20.8	46.2	966	2	US-08-971-217-24	Sequence 24, Appl
C 38	20.8	46.2	966	3	US-09-350-600-24	Sequence 24, Appl
C 39	20.8	46.2	966	3	US-09-906-234-24	Sequence 24, Appl
C 40	20.8	46.2	966	3	US-08-181-271A-21	Sequence 21, Appl
C 41	20.8	46.2	1204	2	US-08-449-315-21	Sequence 21, Appl
C 42	20.8	46.2	1204	2	US-08-444-803-21	Sequence 21, Appl
C 43	20.8	46.2	1204	2	US-08-449-043-21	Sequence 21, Appl
C 44	20.8	46.2	1204	2	US-08-456-265A-21	Sequence 21, Appl
C 45	20.8	46.2	1204	2	US-08-455-416-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-134-000C-3345
Sequence 3345, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3345
LENGTH: 720
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-3345

Query Match 51.6%; Score 23.2; DB 3; Length 720;
Best Local Similarity 77.8%; Pred. No. 7;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGA 37
DB 36 GAAAGATTTCGCGATTCCTTCTTGATTAAGA 71

RESULT 2
US-09-949-016-156329/C
Sequence 156329, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08

```
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 156329
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-156329
```

```
Query Match          51.1%; Score 23; DB 3; Length 601;
Best Local Similarity 74.4%; Pred. No. 8.1;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
OY 5 AGAATTCGCCACCGCTTCTTGATTAAGAATGTTT 43
Db 554 ACAGTTCCCGCATCTCTTCTGATCAAGAGGTTT 516
```

```
RESULT 3
US-09-949-016-16108/c
/ Sequence 16108, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 16108
/ LENGTH: 101300
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(101300)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16108
```

```
Query Match          51.1%; Score 23; DB 3; Length 101300;
Best Local Similarity 74.4%; Pred. No. 29;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
OY 5 AGAATTCGCCACCGCTTCTTGATTAAGAATGTTT 43
Db 36639 ACAGTTCCCGCATCTCTTCTGATCAAGAGGTTT 36601
```

```
RESULT 4
US-09-513-999C-26046
/ Sequence 26046, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclet, A.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ PATENT REFERENCE: 59,US2,REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/122,487
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 26046
```

```
/ LENGTH: 364
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 172
/ OTHER INFORMATION: k=g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 178
/ OTHER INFORMATION: r=a or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 179
/ OTHER INFORMATION: k=g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 181
/ OTHER INFORMATION: r=a or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 182
/ OTHER INFORMATION: s=g or c
US-09-513-999C-26046
```

```
Query Match          50.2%; Score 22.6; DB 3; Length 364;
Best Local Similarity 75.7%; Pred. No. 10;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
OY 9 TTCCGCACCGCTTCTTGATTAAGAATGTTTC 45
Db 88 TTCCCGAGCTTCTTCTTGAGCAAGAATGATTC 124
```

```
RESULT 5
US-09-265-108-1/c
/ Sequence 1, Application US/09265108
/ Patent No. 6033891
/ GENERAL INFORMATION:
/ APPLICANT: Gollightly, Elizabeth
/ APPLICANT: Brown, Kimberly
/ TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides
/ TITLE OF INVENTION: Having Cellobiose Dehydrogenase Activity
/ FILE REFERENCE: 5850,000-US
/ CURRENT APPLICATION NUMBER: US/09/265,108
/ PRIOR FILING DATE: 1999-03-09
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FaastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 2576
/ TYPE: DNA
/ ORGANISM: Humicola
US-09-265-108-1
```

```
Query Match          49.8%; Score 22.4; DB 3; Length 2576;
Best Local Similarity 72.5%; Pred. No. 21;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
OY 1 CGACAGATTCGCCACCGCTTCTTGATTAAGAATG 40
Db 1555 CGAAGATGACGCCACCGCCTCATGAGCTTAACAATG 1516
```

```
RESULT 6
US-09-479-264-1/c
/ Sequence 1, Application US/09479264
/ Patent No. 6280976
/ GENERAL INFORMATION:
/ APPLICANT: Elizabeth J. Gollightly
/ APPLICANT: Kimberly M. Brown
/ TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides
/ TITLE OF INVENTION: Having Cellobiose Dehydrogenase Activity
/ FILE REFERENCE: 5850,200-US
```

CURRENT APPLICATION NUMBER: US/09/479,264
CURRENT FILING DATE: 2000-01-05
EARLIER APPLICATION NUMBER: 09/265,108
EARLIER FILING DATE: 1999-03-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2576
TYPE: DNA
ORGANISM: Humicola
US-09-479-264-1

Query Match 49.8%; Score 22.4; DB 3; Length 2576;
Best Local Similarity 72.5%; Pred. No. 21;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGACAGATTCCCGCCGCTTCTTGATTAAGATG 40
DB 1555 CGAAGATGACGCGCCCTCATTAAGATG 1516

RESULT 7

US-08-976-259-14/c
Sequence 14, Application US/08976259
Patent No. 6316609
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Choi, Gil H.
APPLICANT: Welch, Rodney A.
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609
NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:
ADDRESSER: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,259
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 22671 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-976-259-14

Query Match 48.9%; Score 22; DB 3; Length 22671;
Best Local Similarity 73.7%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCCGACCGCTTCTTGATTAAGATGTTTC 45
DB 4970 ACTCCCAACGCTTCTTGCAAGTTTAAGTTTC 4933

RESULT 8
US-09-956-004-14/c
Sequence 14, Application US/09956004
Patent No. 6787643
GENERAL INFORMATION:
APPLICANT: Patrick J. Dillon et al.
TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
FILE REFERENCE: PB324DI
CURRENT APPLICATION NUMBER: US/09/956,004
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: 08/976,259
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/061,953
PRIOR FILING DATE: 1997-10-14
PRIOR APPLICATION NUMBER: 60/031,626
PRIOR FILING DATE: 1996-11-22
NUMBER OF SEQ ID NOS: 142
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 22671
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc feature
LOCATION: (19750)..(19750)
OTHER INFORMATION: n equals a, t, g, or c
NAME/KEY: misc feature
LOCATION: (20174)..(20174)
OTHER INFORMATION: n equals a, t, g, or c
US-09-956-004-14

Query Match 48.9%; Score 22; DB 3; Length 22671;
Best Local Similarity 73.7%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCCGACCGCTTCTTGATTAAGATGTTTC 45
DB 4970 ACTCCCAACGCTTCTTGCAAGTTTAAGTTTC 4933

RESULT 9
US-09-949-016-12320
Sequence 12320, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12320
LENGTH: 10554
TYPE: DNA
ORGANISM: Human
US-09-949-016-12320

Query Match 48.4%; Score 21.8; DB 3; Length 10554;
Best Local Similarity 78.8%; Pred. No. 52;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 AATTCGACCGCTTCTTGATTAAGAT 39
DB 1500 AATTCACACGCTTCTTAATTAATAAT 1532

RESULT 10
US-09-949-016-16084
; Sequence 16084, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16084
; LENGTH: 10554
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16084

Query Match 48.4%; Score 21.8; DB 3; Length 10554;
Best Local Similarity 78.8%; Pred. No. 52;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 7 AATTCGACCGCTTCTTGATTAAGAT 39
1500 AATTCACCACTTATCTGTAATTAATAAT 1532

RESULT 11
US-09-949-016-15362
; Sequence 15362, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15362
; LENGTH: 51291
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)..(51291)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15362

Query Match 48.4%; Score 21.8; DB 3; Length 51291;
Best Local Similarity 70.7%; Pred. No. 77;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Db 2 GAACAGATTCGACCGCTTCTTGATTAAGATGT 42
38731 GAAGGATTCATTCACATTCATGATTAAGAACTT 38771

RESULT 12
US-09-198-452A-1
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffee, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE REFERENCE: 9710-003-999
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; LOCATION: (1)..(15000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (15001)..(30000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (30001)..(45000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (45001)..(60000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (60001)..(75000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (75001)..(90000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (90001)..(105000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (105001)..(120000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (120001)..(135000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (135001)..(150000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (150001)..(165000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (165001)..(180000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (180001)..(195000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (195001)..(210000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (210001)..(225000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (225001)..(240000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (240001)..(255000)
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a or c or g or t
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t


```
; NAME/KEY: misc_feature
; LOCATION: (270001) ..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001) ..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001) ..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001) ..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001) ..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001) ..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001) ..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001) ..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001) ..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001) ..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001) ..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001) ..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (450001) ..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (465001) ..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (480001) ..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (495001) ..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (510001) ..(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (525001) ..(540000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (540001) ..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (555001) ..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (570001) ..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (585001) ..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (600001) ..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (615001) ..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

```
; LOCATION: (630001) ..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (645001) ..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (660001) ..(675000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (675001) ..(690000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (690001) ..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (705001) ..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (720001) ..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (735001) ..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (750001) ..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (765001) ..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (780001) ..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (795001) ..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (810001) ..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (825001) ..(840000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (840001) ..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (855001) ..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (870001) ..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (885001) ..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (900001) ..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

```
Query Match 48.0%; Score 21.6; DB 3; Length 1230025;
Best Local Similarity 75.0%; Pred No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
Qy 9 TTCCGCCACCGCTTCTGTGATTAGAAATGTTT 44
Db 520397 TTCTGCCATCTCTTGTGACAGCAAGATTGTTT 520432
```

```
RESULT 13
US-09-438-185A-1
; Sequence 1, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
```

```
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1230230
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1

Query Match          48.0%; Score 21.6; DB 3; Length 1230230;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTCCGCCACCGCTTCTGATTAAGATGTTT 44
DB 510309 TTCTGCCATCTCTTGCTAGAGCAAGATGTTT 510344

RESULT 14
US-09-489-039A-20
; Sequence 20, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 20
; LENGTH: 2403
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-20

Query Match          47.6%; Score 21.4; DB 3; Length 2403;
Best Local Similarity 71.8%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 AATTCGCCACCGCTTCTGATTAAGATGTTTC 45
DB 2324 ACTTCGCTTCGCTTCGCTGATGATTCGATTTTTC 2362

RESULT 15
US-09-489-039A-7/C
; Sequence 7, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7
```

```
; LENGTH: 4266
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7

Query Match          47.6%; Score 21.4; DB 3; Length 4266;
Best Local Similarity 71.8%; Pred. No. 61;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 AATTCGCCACCGCTTCTGATTAAGATGTTTC 45
DB 2024 ACTTCGCTTCGCTTCGCTGATGATTCGATTTTTC 1986

Search completed: April 7, 2006, 18:22:39
Job time : 21.0996 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:02:11 ; Search time 73.7527 Seconds
(without alignments)
5045.535 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45

Sequence: 1 CGACGAGATTCGCGACGCGC.....ttgtgtaagaatgttttc 45

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Database :

Published Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/us08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/us09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/us10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/us10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/us10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/us10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/us10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/us11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	45	3	US-09-211-691-4
2	45	100.0	45	6	US-10-317-773-4
3	45	100.0	45	6	US-10-317-773-4
4	27	60.0	623	9	US-10-915-740A-380
5	27	60.0	65632	9	US-10-915-740A-49
6	27	60.0	2242716	9	US-10-915-740A-1068
7	24	53.3	367	8	US-10-357-930-36950
8	24	53.3	367	8	US-10-357-930-42575
9	23	52.4	31	8	US-10-819-861A-9
10	23	52.0	351	3	US-09-867-701-3041
11	22	49.3	3371	6	US-10-108-260A-1181
12	22	49.3	546	7	US-10-021-323-11510
13	22	49.8	1035	7	US-09-925-065A-37059
14	22	49.3	619	5	US-10-027-632-102109
15	22	49.3	619	5	US-10-027-632-102110
16	22	49.3	619	6	US-10-027-632-102110
17	22	49.3	619	6	US-10-027-632-102110
18	22	48.9	551	4	US-09-925-065A-364339
19	22	48.9	2914	7	US-10-437-963-83308
20	22	48.9	6483	3	US-09-880-107-3699
21	22	48.9	6483	3	US-10-335-696-24
22	22	48.9	22671	3	US-09-956-004-14
23	22	48.9	22671	8	US-10-808-570-14

24	21.8	48.4	412	4	US-09-925-065A-134253	Sequence 134253,
25	21.8	48.4	414	4	US-09-925-065A-179113	Sequence 179113,
26	21.8	48.4	461	5	US-10-027-632-71563	Sequence 71563, A
27	21.8	48.4	461	5	US-10-027-632-71564	Sequence 71564, A
28	21.8	48.4	461	5	US-10-027-632-71565	Sequence 71565, A
29	21.8	48.4	461	5	US-10-027-632-108559	Sequence 108559,
30	21.8	48.4	461	5	US-10-027-632-108560	Sequence 108560,
31	21.8	48.4	461	5	US-10-027-632-108561	Sequence 108561,
32	21.8	48.4	461	5	US-10-027-632-180103	Sequence 180103,
33	21.8	48.4	461	5	US-10-027-632-312469	Sequence 312469,
34	21.8	48.4	461	5	US-10-027-632-312470	Sequence 312470,
35	21.8	48.4	461	5	US-10-027-632-312471	Sequence 312471,
36	21.8	48.4	461	6	US-10-027-632-71563	Sequence 71563, A
37	21.8	48.4	461	6	US-10-027-632-71564	Sequence 71564, A
38	21.8	48.4	461	6	US-10-027-632-71565	Sequence 71565, A
39	21.8	48.4	461	6	US-10-027-632-108559	Sequence 108559,
40	21.8	48.4	461	6	US-10-027-632-108560	Sequence 108560,
41	21.8	48.4	461	6	US-10-027-632-108561	Sequence 108561,
42	21.8	48.4	461	6	US-10-027-632-180103	Sequence 180103,
43	21.8	48.4	461	6	US-10-027-632-312469	Sequence 312469,
44	21.8	48.4	461	6	US-10-027-632-312470	Sequence 312470,
45	21.8	48.4	461	6	US-10-027-632-312471	Sequence 312471,

ALIGNMENTS

RESULT 1
US-09-211-691-4
; Sequence 4, Application US/09211691
; Patent No. US20020034805A1
GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
; FILE REFERENCE: 019957-012910US
; CURRENT APPLICATION NUMBER: US/09/211,691
; CURRENT FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:SYNTH-R6 3'
US-09-211-691-4
Query Match
Best Local Similarity 100.0%; Score 45; DB 3; Length 45;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACGAGATTCGCGACGCGCTTCTTGATTAAGAAATGTTTC 45
DB 1 CGACGAGATTCGCGACGCGCTTCTTGATTAAGAAATGTTTC 45
RESULT 2
US-10-317-773-4
; Sequence 4, Application US/10317773
; Publication No. US20030180928A1
GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNAc 4' Epimerase and a

```
/ TITLE OF INVENTION: GalNac Transferase
/ FILE REFERENCE: 019633-00081205
/ CURRENT APPLICATION NUMBER: US/10/317,773
/ CURRENT FILING DATE: 2002-12-11
/ PRIOR APPLICATION NUMBER: US 60/069,443
/ PRIOR FILING DATE: 1997-12-15
/ PRIOR APPLICATION NUMBER: US 09/211,691
/ PRIOR FILING DATE: 1998-12-14
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:SYNTH-R6 3'
US-10-317-773-4
```

```
Query Match          100.0%; Score 45; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
DB 1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
```

```
RESULT 3
US-10-317-428-4
/ Sequence 4, Application US/10317428
/ Publication No. US20030186414A1
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
/ APPLICANT: Young, N. Martin
/ APPLICANT: Wakarchuk, Warren W.
/ APPLICANT: National Research Council of Canada
/ TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
/ FILE REFERENCE: 019633-00081105
/ CURRENT APPLICATION NUMBER: US/10/317,428
/ CURRENT FILING DATE: 2002-12-11
/ PRIOR APPLICATION NUMBER: US 60/069,443
/ PRIOR FILING DATE: 1997-12-15
/ PRIOR APPLICATION NUMBER: US 09/211,691
/ PRIOR FILING DATE: 1998-12-14
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 45
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:SYNTH-R6 3'
US-10-317-428-4
```

```
Query Match          100.0%; Score 45; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
DB 1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
```

```
RESULT 4
US-10-915-740A-380
/ Sequence 380, Application US/10915740A
/ Publication No. US20050191316A1
/ GENERAL INFORMATION:
/ APPLICANT: Frazer, Claire M.
/ APPLICANT: Hickey, Erin
/ APPLICANT: Peterson, Jeremy
```

```
/ APPLICANT: Tettelein, Herve
/ APPLICANT: Venter, J. Craig
/ APPLICANT: Massignani, Vega
/ APPLICANT: Galeotti, Cesira
/ APPLICANT: Mora, Manroza
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Scarselli, Maria
/ APPLICANT: Scarlato, Vincenzo
/ APPLICANT: Rappuoli, Rino
/ APPLICANT: Pizza, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
/ FILE REFERENCE: 002441.00090
/ CURRENT APPLICATION NUMBER: US/10/915,740A
/ CURRENT FILING DATE: 2004-08-11
/ PRIOR APPLICATION NUMBER: 09/806,866
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: USSN 60/103,794
/ PRIOR FILING DATE: 1998-10-09
/ PRIOR APPLICATION NUMBER: USSN 60/132,068
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/US99/25373
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 1068
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 380
/ LENGTH: 623
/ TYPE: DNA
/ ORGANISM: Neisseria meningitidis
US-10-915-740A-380
```

```
Query Match          60.0%; Score 27; DB 9; Length 623;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 19 GCTTTCCTTGATTAAGAAGTTTTC 45
DB 161 GCTTTCCTTGATTAAGAAGTTTTC 187
```

```
RESULT 5
US-10-915-740A-49/C
/ Sequence 49, Application US/10915740A
/ Publication No. US20050191316A1
/ GENERAL INFORMATION:
/ APPLICANT: Frazer, Claire M.
/ APPLICANT: Hickey, Erin
/ APPLICANT: Peterson, Jeremy
/ APPLICANT: Tettelein, Herve
/ APPLICANT: Venter, J. Craig
/ APPLICANT: Massignani, Vega
/ APPLICANT: Galeotti, Cesira
/ APPLICANT: Mora, Manroza
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Scarselli, Maria
/ APPLICANT: Scarlato, Vincenzo
/ APPLICANT: Rappuoli, Rino
/ APPLICANT: Pizza, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
/ FILE REFERENCE: 002441.00090
/ CURRENT APPLICATION NUMBER: US/10/915,740A
/ CURRENT FILING DATE: 2004-08-11
/ PRIOR APPLICATION NUMBER: 09/806,866
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: USSN 60/103,794
/ PRIOR FILING DATE: 1998-10-09
/ PRIOR APPLICATION NUMBER: USSN 60/132,068
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/US99/25373
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 1068
/ SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 49
; LENGTH: 65632
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-49

Query Match      60.0%; Score 27; DB 9; Length 65632;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GCTTTCCTTGATTAAGATGTTTC 45
Db      55738 GCTTTCCTTGATTAAGATGTTTC 55712

RESULT 6
US-10-915-740A-1068
; Sequence 1068, Application US/10915740A
; Publication No. US20050191316A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteijn, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Maignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scariato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizzi, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; PRIOR FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1068
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-1068

Query Match      60.0%; Score 27; DB 9; Length 2242716;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GCTTTCCTTGATTAAGATGTTTC 45
Db      78420 GCTTTCCTTGATTAAGATGTTTC 78446

RESULT 7
US-10-357-930-36950
; Sequence 36950, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
```

```
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36950
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 340..341
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-36950

Query Match      53.3%; Score 24; DB 8; Length 367;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy      4 CAGATTCCGACCGCTTCTTGATTAAGATGTT 43
Db      271 CAGATTAGCCTTACTTCTTGATTAAGATGTT 310

RESULT 8
US-10-357-930-42575
; Sequence 42575, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42575
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: misc.feature
LOCATION: 340_341
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-42575

Query Match
Best Local Similarity 53.3%; Score 24; DB 8; Length 367;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CAGAAATCCGCCACCGCTTCTCTGTGATTAAGATGTTT 43
Db 271 CACAAATTAAGCTTACTTCTCTGTGATTAAGAAATGTTT 310

RESULT 9
US-10-819-861A-9
Sequence 9, Application US/10819861A
Publication No. US20040241190A1
GENERAL INFORMATION:
APPLICANT: Xenova Research Limited
APPLICANT: McLean, Cornelia S.
APPLICANT: Keen, Simon W.
APPLICANT: Martin, Gillian May
TITLE OF INVENTION: Vaccine Preparations
FILE REFERENCE: 5673-68171-01
CURRENT APPLICATION NUMBER: US/10/819,861A
CURRENT FILING DATE: 2004-04-06
PRIOR APPLICATION NUMBER: GB 0308691.5
PRIOR FILING DATE: 2003-04-07
PRIOR APPLICATION NUMBER: US 60/464,758
PRIOR FILING DATE: 2003-04-21
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9
LENGTH: 31
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide primer.
US-10-819-861A-9

Query Match
Best Local Similarity 52.4%; Score 23.6; DB 8; Length 31;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 10 TCCGCCACCGCTTCTCTGTGATTAAGAT 39
Db 2 TCGATATCGCTTCTCTGTGATTAAGAT 31

RESULT 10
US-09-867-701-3041/C
Sequence 3041, Application US/09867701
Patent No. US2002013237A1
GENERAL INFORMATION:
APPLICANT: Aglate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Hailocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.497
CURRENT APPLICATION NUMBER: US/09/867,701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 3041
LENGTH: 351
TYPE: DNA
ORGANISM: Homo sapien
US-09-867-701-3041

Query Match 52.0%; Score 23.4; DB 3; Length 351;

Best Local Similarity 73.2%; Pred. No. 23;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GACAGAAATCCGCCACCGCTTCTCTGTGATTAAGATGTTT 42
Db 240 GACAGAAATCCGCCACCGCTTCTCTGTGATTAAGATGTTT 200

RESULT 11
US-10-108-260A-1181
Sequence 1181, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1el full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1181
LENGTH: 3371
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-1181

Query Match 50.7%; Score 22.8; DB 6; Length 3371;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 3 ACAGAAATCCGCCACCGCTTCTCTGTGATTAAGATGTTT 44
Db 1378 ACAGAAATCCGCCACCGCTTCTCTGTGATTAAGATGTTT 1419

RESULT 12
US-10-021-323-11510
Sequence 11510, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255,619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 11510
LENGTH: 546
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3829-002-01-K6-H1
US-10-021-323-11510

Query Match 49.8%; Score 22.4; DB 7; Length 546;
Best Local Similarity 72.5%; Pred. No. 69;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 5 AGAATTCGCCACCGCTTCTCTGTGATTAAGATGTTT 44
Db 80 AGAATTCGCCACCGCTTCTCTGTGATTAAGATGTTT 119

RESULT 13
US-09-925-065A-37059
Sequence 37059, Application US/09925065A
Publication No. US20050228172A9

```
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 37059
/ LENGTH: 1035
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-925-065A-37059

Query Match
Best Local Similarity 49.8%; Score 22.4; DB 4; Length 1035;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 CAGAAATCCGCCACCGCTTCTTGATTAAGAAATTT 43
DB 871 CCGAATTCCTCCGAGATTCCCTTGATTAATCTATT 910

RESULT 14
US-10-027-632-102109
/ Sequence 102109, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 102109
/ LENGTH: 619
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(619)
/ OTHER INFORMATION: n = A,T,C or G
/ US-10-027-632-102109

Query Match
Best Local Similarity 49.3%; Score 22.2; DB 5; Length 619;
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
```

```
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 ACAGAAATCCGCCACCGCTTCTTGATTAAGAAAT 39
DB 114 AAAGGCTAACGCTCCACCTTCTTGATTAAGAAAT 150

RESULT 15
US-10-027-632-102110
/ Sequence 102110, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 102110
/ LENGTH: 619
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(619)
/ OTHER INFORMATION: n = A,T,C or G
/ US-10-027-632-102110

Query Match
Best Local Similarity 49.3%; Score 22.2; DB 5; Length 619;
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 ACAGAAATCCGCCACCGCTTCTTGATTAAGAAAT 39
DB 114 AAAGGCTAACGCTCCACCTTCTTGATTAAGAAAT 150

Search completed: April 7, 2006, 18:56:34
Job time : 76.7527 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:05:06 ; Search time 52.5328 Seconds
(without alignments)
3426.906 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45

Sequence: 1 CGACAGATTCGCCCGCCGCGC.....ttgtgataagaatgttttc 45

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA_New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
10: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23.4	52.0	159497	14 US-11-112-908-61	Sequence 61, Appl
C 2	23.4	52.0	171427	14 US-11-112-908-60	Sequence 60, Appl
C 3	22.4	49.8	1035	US-09-925-065A-37059	Sequence 37059, A
4	22.4	49.8	1035	US-10-301-480-138297	Sequence 138297,
5	22.4	49.8	1035	US-10-301-480-751706	Sequence 751706,
C 6	22	48.9	544	US-10-301-480-434999	Sequence 434999,
C 7	22	48.9	544	US-10-301-480-104608	Sequence 104608,
C 8	22	48.9	551	US-09-925-065A-364339	Sequence 364339,
C 9	21.8	48.4	200	US-11-098-686-7073	Sequence 7073, Ap
10	21.8	48.4	412	US-09-925-065A-134253	Sequence 134253,
11	21.8	48.4	412	US-10-301-480-10543	Sequence 10543, A
12	21.8	48.4	412	US-10-301-480-623952	Sequence 623952,
13	21.8	48.4	414	US-09-925-065A-179113	Sequence 179113,
14	21.8	48.4	419	US-10-301-480-229997	Sequence 229997,
15	21.8	48.4	419	US-10-301-480-843406	Sequence 843406,
16	21.8	48.4	435	US-10-301-480-10542	Sequence 10542, A
17	21.8	48.4	435	US-10-301-480-623951	Sequence 623951,
18	21.8	48.4	581	US-10-301-480-333515	Sequence 333515,

19	21.8	48.4	581	10 US-10-301-480-346924	Sequence 346924,
20	21.8	48.4	588	6 US-09-925-065A-254602	Sequence 254602,
21	21.8	48.4	588	6 US-09-925-065A-254603	Sequence 254603,
22	21.8	48.4	588	10 US-10-301-480-333516	Sequence 333516,
23	21.8	48.4	588	10 US-10-301-480-346925	Sequence 346925,
24	21.8	48.4	616	6 US-09-925-065A-854928	Sequence 854928,
25	21.8	48.4	1000	10 US-10-301-480-580482	Sequence 580482,
26	21.8	48.4	1000	10 US-10-301-480-1193891	Sequence 1193891,
27	21.8	48.4	4167	14 US-11-098-686-9804	Sequence 9804, Ap
C 28	21.8	48.4	145619	14 US-11-098-686-8739	Sequence 8739, Ap
C 29	21.6	48.0	571	6 US-09-925-065A-195299	Sequence 195299,
C 30	21.6	48.0	580	10 US-10-301-480-284500	Sequence 284500,
C 31	21.6	48.0	580	10 US-10-301-480-897909	Sequence 897909,
C 32	21.6	48.0	837	6 US-09-925-065A-678556	Sequence 678556,
C 33	21.6	48.0	1380	9 US-10-301-480-25075	Sequence 25075, A
C 34	21.6	48.0	1380	10 US-10-301-480-638484	Sequence 638484,
35	21.6	48.0	37415	9 US-10-330-773-486	Sequence 486, App
36	21.4	47.6	457	6 US-09-925-065A-443892	Sequence 443892,
37	21.4	47.6	457	10 US-10-301-480-504230	Sequence 504230,
38	21.4	47.6	457	10 US-10-301-480-1117639	Sequence 1117639,
39	21.4	47.6	575	6 US-09-925-065A-391550	Sequence 391550,
40	21.4	47.6	585	10 US-10-301-480-459770	Sequence 459770,
41	21.4	47.6	585	10 US-10-301-480-1073179	Sequence 1073179,
42	21.4	47.6	994	10 US-10-301-480-566796	Sequence 566796,
43	21.4	47.6	994	10 US-10-301-480-1180205	Sequence 1180205,
44	21.2	47.1	819	6 US-09-925-065A-4742	Sequence 4742, Ap
45	21.2	47.1	819	6 US-09-925-065A-4743	Sequence 4743, Ap

ALIGNMENTS

```
RESULT 1
US-11-112-908-61/c
; Sequence 61, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 159497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-61

Query Match
Best Local Similarity 73.2%; Pred. No. 76;
Matches 30; Conservative 0; Mismatches 11; Indels 0;
Gaps 0;

Cy 2 GACGAGATTCGCCCGCCGCTTCTGATTAAGATGTT 42
Db 127197 GACGAGATTCGCCCGCCGCTTCTGATTAAGATGTT 127157

RESULT 2
US-11-112-908-60/c
; Sequence 60, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60
; LENGTH: 171427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-60

Query Match      52.0%; Score 23.4; DB 14; Length 171427;
Best Local Similarity 73.2%; Pred. No. 78;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      2 GACAGATTCCGCCACCGCTTCTTGATTAAGAATGTT 42
Db      33333 GACAGATTCCGCCACCGCTTCTTGATTAAGAATGTT 33293
```

```

RESULT 3
US-09-925-065A-37059
; Sequence 37059, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37059
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-37059
```

```

Query Match      49.8%; Score 22.4; DB 6; Length 1035;
Best Local Similarity 72.5%; Pred. No. 46;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      4 CAGAAATCCGCCACCGCTTCTTGATTAAGAATGTT 43
Db      871 CTGAATCTTCCAGAGTTTCCCTTGATTAATCTAATT 910
```

```

RESULT 4
US-10-301-480-138297
; Sequence 138297, Application US/10301480
; Publication No. US20060057564A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138297
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-138297
```

```

Query Match      49.8%; Score 22.4; DB 9; Length 1035;
Best Local Similarity 72.5%; Pred. No. 46;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      4 CAGAAATCCGCCACCGCTTCTTGATTAAGAATGTT 43
Db      871 CTGAATCTTCCAGAGTTTCCCTTGATTAATCTAATT 910
```

```

RESULT 5
US-10-301-480-751706
; Sequence 751706, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 751706
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-751706
```

```

Query Match      49.8%; Score 22.4; DB 10; Length 1035;
Best Local Similarity 72.5%; Pred. No. 46;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      4 CAGAAATCCGCCACCGCTTCTTGATTAAGAATGTT 43
Db      871 CTGAATCTTCCAGAGTTTCCCTTGATTAATCTAATT 910
```

```

RESULT 6
US-10-301-480-434999/C
; Sequence 434999, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
```

```

1 PRIOR APPLICATION NUMBER: US 60/311,695
2
3 PRIOR FILING DATE: 2001-08-10
4
5 NUMBER OF SEQ ID NOS: 122618
6
7 SOFTWARE: PatSeq for Windows Version 4.0.0
8
9 SEQ ID NO 434999
10
11 LENGTH: 544
12
13 TYPE: DNA
14
15 ORGANISM: Homo sapien
16
17 US-10-301-480-434999

```

Query Match	48.9%	Score 22;	DB 10;	Length 544;
Best Local Similarity	73.7%	Pred. No. 55;		
Matches 28; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

QY 8 ATTCCGCACCGCTTTCCTGTGATTAGATGTTTC 45
 ||||| ||||| ||||| ||||| |||||
Db 508 ATTTCCCCCTTGCTTCTGTGATAAGAATGATTTTC 471

```

RESULT 7
US-10-301-480-10484b8/c
; Sequence 1048408, Application US/10301480
; Publication No. US2006057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1048408
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-1048408

```

Query Match	48.9%	Score 22;	DB 10;	Length 544;
Best Local Similarity	73.7%;	Pred. NO. 55;		
Matches 28; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

Qy 8 ATTCCGCCACCGCTTTCCTTGATTAAGAAATGTTTTC 45
Db 508 ATTTCCCCCTTGTCTCTTGTGATAATGAATGATTTTC 471

```

RESULT 8
US-09-925-065A-364339/c
Sequence 364339, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms In the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIORITY APPLICATION NUMBER: US 60/243, 096
PRIORITY FILING DATE: 2000-10-24
PRIORITY APPLICATION NUMBER: US 60/255, 147
PRIORITY FILING DATE: 2000-11-20
PRIORITY APPLICATION NUMBER: US 60/250, 092
PRIORITY FILING DATE: 2000-11-30
PRIORITY APPLICATION NUMBER: US 60/261, 766
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: US 60/289, 846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086

```

```

; SOFTWARE: FASTSQ for Windows Version 4.0
; SEQ ID NO 364339
;
; LENGTH: 551
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
US-09-925-0654-364339

```

Query Match	48.9%	Score 22;	DB 6;	Length 551;
Best Local Similarity	73.7%	Pred. No. 55;		
Matches 28; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

QY 8 ATTCGGCACCGCTTCCGTGATTAAGATGTTTC 45
 ||| ||| ||| ||| ||| ||| ||| |||
Db 508 ATTCCCCCTTGCTTCTTGTAATGAATGATTTTC 471

```

RESULT 9
US-11-098-686-7073/c
; Sequence 7073, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gabhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/099,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7073
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-7073

```

Query Match	48.4%;	Score 21.8;	DB 14;	Length 200;
Best Local Similarity	70.7%;	Pred. No. 50;		
Matches 29; Conservative	0;	Mismatches 12;	Indels 0;	Gaps 0;

```
QY      4 CAGATTCCGCCACCGCTTTCCTTGATTAAGAATGTTTT   44  
        ||||| | | | | | | | | | | | | | |  
Db     151 CAGAATTACGTCCCCTGTTCCTTTGATGGAGCAGTTTT   111
```

```

RESULT 10
US-09-925-065A-134253
Sequence 134253, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIORITY FILING DATE: 2001-08-08
PRIORITY APPLICATION NUMBER: US 60/243,096
PRIORITY FILING DATE: 2000-10-24
PRIORITY APPLICATION NUMBER: US 60/252,147
PRIORITY FILING DATE: 2000-11-20
PRIORITY APPLICATION NUMBER: US 60/250,092
PRIORITY FILING DATE: 2000-11-30
PRIORITY APPLICATION NUMBER: US 60/261,766
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: US 60/289,846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134253
LENGTH: 412

```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-134253

Query Match 48.4%; Score 21.8; DB 6; Length 412;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAAGGT 361

RESULT 11
US-10-301-480-10543
Sequence 10543, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT FILING DATE: 2002-11-21
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10543
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-10543

Query Match 48.4%; Score 21.8; DB 9; Length 412;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAAGGT 361

RESULT 12
US-10-301-480-623952
Sequence 623952, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT FILING DATE: 2002-11-21
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 623952
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-623952

Query Match 48.4%; Score 21.8; DB 10; Length 412;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42

DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAAGGT 361

RESULT 13
US-09-925-065A-179113
Sequence 179113, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT FILING DATE: 2001-08-08
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 179113
LENGTH: 414
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-179113

Query Match 48.4%; Score 21.8; DB 6; Length 414;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAAGGT 361

RESULT 14
US-10-301-480-229997
Sequence 229997, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT FILING DATE: 2002-11-21
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 229997
LENGTH: 419
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-229997

Query Match 48.4%; Score 21.8; DB 10; Length 419;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAAGGT 361

```

RESULT 15
US-10-301-480-843406
; Sequence 843406, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 843406
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-843406
    
```

```

Query Match      48.4%; Score 21.8; DB 10; Length 419;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      2  GACGAGATTCGCGACCGCTTCTCTTGATTAAGATGTT 42
Db      321 GCCGCGATTCGCGAAGCTTACTTTGATTAATGAAGGT 361
    
```

Search completed: April 7, 2006, 21:23:53
 Job time : 55.5328 secs

THIS PAGE BLANK (USPTO)